تطبيقات إحصائية في بحوث الإعلام

إعداد

الدكتور/ بركات عبد العزيز

أستاذ بكلية الإعلام – جامعة القاهرة

المحتويات

الفصل الأول: مدخل تمهيدي

الفصل الأول

مدخل تمهيدي

أولاً: الإحصاء وأهميته للعلم:

الإحصاء في اللغة يعني العد الشامل، ومن المجاز قول العرب لم أر أكثر منهم حصى، أي لم أر أكثر منهم عدداً، وقولهم هذا أمر لا أحصيه أي لا أطيقه ولا أضبطه. و الإحصاء كما يفهمه عامة الناس لا يخرج عن كونه جمع معلومات رقمية وعرضها في جداول ورسوم بيانية، كما يفهمه البعض في إطار حساب المتوسطات والنسب المختلفة، لكن الإحصاء في حقيقته وأساليبه المتطورة يتجاوز هذه النظرة بمراحل هائلة. وقد نشأ علم الإحصاء في إطار التنظيم السياسي للدولة على يد البارون بيلفد F. Von Biefe'd سنة ١٧٧٠ ، غير أنه سرعان ما تطورت الأصول الرياضية لعلم الإحصاء بواسطة علماء الرياضيات الأوروبيين مثل لابلاس Laplace الرياضي الفرنسي وجاوس Gaus الرياضي الألماني، وجولتان Galton العالم الانجليزي وكارل بيرسون Karl Person الرياضي الانجليزي، وغيرهم من العلماء الذين تواصلت جهودهم العلمية حتى وصل الإحصاء إلى ما هو عليه الآن كعلم متطور يرتبط بكافة مجالات الحياة

والإحصاء في صورته الحديثة هو إحدى الدعامات الرئيسية التي تقوم عليها طرق البحث العلمي في التخصصات المعرفية المختلفة، فهذه الطرق في جوهرها العام لا تخرج عن إجراء تجارب ميدانية، ثم استخلاص النتائج الموضوعية المستمدة من تلك التجارب، وصياغة القوانين والنظريات التي تفسر تلك النتائج. ويرتبط الإحصاء ارتباطاً وثيقاً بتلك الخطوات، حيث يحدد الشروط الأساسية لموضوعية التجارب ومنهجيتها، كما يحدد طرق التحليل المناسبة لكل تجربة ومدى التعميم الذي تنطوي عليه نتائج تلك التجارب... بهذا المعنى فإن الإحصاء يشكل الأساس القاعدي لضبط تنفيذ الأبحاث الحديثة في العلوم المعاصرة التي تقوم على الملاحظة الدقيقة والتجريب العلمي والتحليل الرياضي والاستنتاج المنطقي، وبهذه الطريقة تتخذ العلوم صفة التجريبية والموضوعية.

وبموجب الطريقة العلمية يتم جمع البيانات عن الظواهر المختلفة ثم معالجة تلك البيانات باستخدام التحليل الإحصائي بما يمكن من رصد النتائج رصداً موجزاً واضحاً، ومن ثم معرفة خصائص الظاهرة وأبعادها وما تنطوي عليه من دلالات، والهدف من ذلك فهم

العوامل الأساسية التي تؤثر في الظاهرة محل البحث، وقد يتم التوصل من هذا كله إلى الكشف عن الفكرة الجوهرية أو القانون العام الذي يصلح لتفسير تلك الظاهرة والظواهر الأخرى ذات العلاقة بها، ومن هنا كان الإحصاء من أهم الوسائل التي تتم الاستعانة بها في العلوم المختلفة في الوصول إلى النتائج الهامة وتطبيقها.

وإذا كان العلم في جوهره تنظيم موضوعي للفكر يقوم على تبادل المعرفة بين المشتغلين بالبحث، فإن أغلب الأبحاث الحديثة تعتمد على الأرقام والمعالجة الإحصائية للبيانات المختلفة، ولهذا يتعين على المعنيين بالبحث العلمي أن يعرفوا أدواته ليسايروا تطوره ومناهجه وتطبيقاته المتنوعة، فالتطور العلمي لأي فرع من فروع المعرفة يقاس بمدى تطور هذه الجوانب.وقد أحرزت العلوم الطبيعية قصب السبق في هذا المضمار لبساطة تكوينها وثبوت نتائجها وإمكانية إخضاعها للضبط العلمي الدقيق، كما أن تلك العلوم استعانت مبكراً بالأعداد والعلوم الرياضية، أما العلوم الاجتماعية والإنسانية فقد تأخرت في نشأتها الأولي عن هذا التطور لأسباب كثيرة تتعلق بطبيعة تلك العلوم من حيث التعقيد واتساع مجالاتها وصعوبة إخضاعها للتجريب والضبط العلمي الكميQuantitative لكن العلوم الاجتماعية والإنسانية سرعان ما تطورت من حيث الطرق والمنهج وأدمجت الضبط الكمي في دراسة الظواهر المختلفة، وأصبحت تستخدم الكثير من الطرق الكمية المتقدمة بالارتكاز على علم الإحصاء وتطبيقاته

ثانياً:الإحصاء وخطوات البحث العلمي

يرتبط الإحصاء ارتباطاً وثيقاً بالبحث العلمي الذي يتسارع في عصرنا الحاضر، وينتظم البحث العلمي في خطوات منطقية واضحة، ويستقيم البحث العلمي عندما يعتمد على التحليل الإحصائي في خطواته المتفاعلة، والتي تتمثل في اختيار المشكلة، وضع التساؤلات والفروض، تصميم خطة البحث، جمع البيانات ومعالجتها وتنظيمها وتفسير النتائج وصياغتها في تقرير علمي متكامل..... في كل هذه الخطوات يبدو الإحصاء ذا أهمية، وفيما يلي نبذة موجزة عن تلك الخطوات:

(١) اختيار المشكلة:

يبدأ البحث بمشكلة عامة تتطور خلال التحليل إلى مشكلة محددة تتطلب إجابات محددة، ومن المعروف أن اختيار المشكلة وصياغتها صياغة دقيقة هي التي تجعلها قابلة للبحث، وتتلخص أهم الأسس الرئيسية لاختيار المشكلة في:

- ألا تكون كبيرة واسعة حتى لا يتم بحثها بضحالة أو يصعب بحثها، وألا تكون ضيقة جداً أو محدودة حتى لا تصبح تافهة، بل تكون وسطاً بين هذه وتلك حتى يمكن دراستها في يسر والتوصل إلى نتائج ذات قيمة علمية
- أن يكون في الإمكان بحثها خلال المدة الزمنية المقررة، وأن تكون تكلفتها في حدود الإمكانات المتاحة
- أن تكون جديدة بحيث تكشف عن بعض الآفاق المجهولة، وإلا فقدت قوتها وأهميتها
- أن تتفق وميل الباحث ومستوى قدرته على معالجتها، وأن تكون بياناتها المختلفة ميسورة بحيث لا تكلف الباحث عنتاً أو مشقة بالغة

(٢) صياغة التساؤلات والفروض:

بعد تحديد المشكلة البحثية بدقة، يقوم الباحث بطرح التساؤلات التي تعبر عن أبعاد المشكلة، وهذه التساؤلات هي التي يستهدف البحث أن يجيب عليها حتى يتم حسم المشكلة البحثية من جوانبها المختلفة, وتقتصر بعض البحوث على طرح مجموعة من التساؤلات وذلك عندما يكون التخصص يفتقد التراكم المعرفي المناسب، أو في حالة ندرة الدراسات السابقة حول الموضوع، أما إذا توافرت تلك الدراسات أو كان مجال التخصص العلمي من العمق والثراء الكافيين،فإن الباحث يضع التساؤلات كتعبير عن المشكلة البحثية بأبعادها المختلفة، ثم يصيغ مجموعة فروض بحيث يكون كل فرض بمثابة إجابة محتملة عن تساؤل معين، فعلاقة الفرض بالمشكلة البحثية هي علاقة الإجابة بالسؤال الذي يعبر عن أحد جوانب المشكلة. والفروض بهذا المعني هي ملتقى الطرق التي

تنتهي إليها المشكلة ويبدأ معها البحث، أي أن موقع الفروض من خطوات البحث يمثـل نقطة التحول من بناء البحث إلي التصميم التجريبي للإجابة على المشكلة القائمة. والحكم الذي يقرر قبول الفرض أو رفضه هو النتيجة التي تنتهي إليها جميع خطـوات البحـث، والتوصل لمثل هذا الحكم يقتضي إجراء التجارب التي تختبر صحة تلك الفروض. وبما أن الطريقة التي يصاغ بها الفرض تؤثر تأثيراً مباشراً على البناء التجريبي للبحث وعلى أدوات جمع البيانات والمعالجة الإحصائية- فإن أي تعقيد أو خطأ في صياغة الفرض يـؤدي إلي مشكلات في تنفيذ البحث، وقد تحول أخطاء الصياغة دون إنجاز البحث، لذلك يجب أن تخضع الفروض لشروط عملية دقيقة نلخص أهمها فيما يلي:

(أ) وحدة الإجابة: يجب أن يكون الفرض في صورة إجابة محددة علـى جانـب محدد من المشكلة التي يتناولها البحث. وليس معنى هذا أن يقتصر البحث على فرض واحد، بل إن الفروض تتعدد بتعدد أبعاد المشكلة البحثية، وبذلك يصبح كل بعد من أبعاد المشكلة الأساسية مشكلة فرعية يعبر عنها فـرض واحـد، والفروض التي تتصدي للإجابة على أكثر من مشكلة فرعية تتطلب بناءً تجريبياً معقداً، ولذلك يجب الاقتصاد في الفروض بحيث لا تكون من الكثـرة الـتي تؤدي إلى التداخل أو التكرار، وقد تعرقل إنجاز البحث.

(ب) البساطة: يجب أن يكون الفرض أبسط إجابة للمشكلة، وكلما كان الفرض بسيطاً مباشراً كان تصميم البحث قابلاً للتنفيذ، علماً بأن الفرض المركب يؤدي إلي بناء تجريبي معقد.

(ج) إمكانية الاختبار: بمعنى أن يكون الفرض متضمناً متغيرات يمكن قياسـها، وعلاقات يمكن التحقق منها، وأن يكون في الإمكان قبـول الفـرض أو عدم قبوله.فإذا كانت صياغة الفرض تؤدي إلي قبوله ولا تؤدي إلي رفضه، أو تؤدي إلى رفضه ولا تؤدي إلى قبوله، فإنه لا يصلح أن يكون فرضاً من فروض البحث. فمثلا الفرض الذي يقول: "إن النـاس يقـاتلون لأن لديهم نزعات عدوانية"

هذا الفرض يمكن قبوله ولا يمكن رفضه لأن قبوله يقتضي ظهور الترعات العدوانية، كما أن رفضه يقتضي اختفاء التراعات العدوانية، والاختفاء المؤقت لهذه التراعات لا يعني عدم وجودها، فقد تكون تلك التراعات كامنة لا تظهر إلا عندما تستشار.

(٣) تصميم خطة البحث:

تقوم خطة البحث على بناء علمي متماسك يسبق القيام بالبحث، وقد تشتمل هذه الخطة على نموذج مصغر للبحث وذلك للكشف عن نواحي قوته وضعفه، والتغلب على الصعوبات التي قد تواجهه، ولتبيان أوضح المسالك لمعالجة المشكلة معالجة علمية دقيقة. فالخطة بهذا المعني تشبه النموذج المصغر أو الرسم التوضيحي للبحث، ومن الضروري أن تشتمل خطة دراسة المشكلة على المعلومات النظرية وأدوات جمع البيانات الميدانية، وعينة الأفراد التي تستخدم في التجربة والأسس العلمية لاختيارها..... الخ

ويمكن اختبار خطة الدراسة من خلال إجراء تجربة تمهيدية على نطاق صغير للكشف عن أثر الظروف المختلفة في نتائج التجربة وتعرف العوائق التي يمكن أن تعترض التنفيذ، أو التي قد تعوق نمو البحث، وكذلك الكشف عن الأخطاء والغموض وأوجه النقص..... مع وضع الأساليب الكفيلة بالتغلب على الصعوبات، وضمان تنفيذ البحث بدقة بحيث يتطور البحث عبر خطوات التنفيذ حتى يصل إلى هدفه النهائي.

(٤) جمع البيانات :

هذه المرحلة هي جوهر التنفيذ العملي لخطة البحث، حيث يتم جمع البيانات النظرية من خلال المصادر المكتبية وأوعية المعلومات التقليدية والحديثة، أما البيانات الميدانية فيتم جمعها باستخدام الأدوات المناسبة بما يتفق وطبيعة البحث، ومن أبرز تلك الأدوات: الاستبانات، الاختبارات، المقاييس، جداول الملاحظة، صحيفة تحليل المحتوى، المقابلات المفتوحة، مجموعات النقاش المركزة....)، ويتم اختيار أو تصميم أدوات جمع البيانات التي تناسب البحث، بحيث يمكن الحصول على المعلومات التي تجيب على التساؤلات أو تتحقق من الفروض

(٥) معالجة البيانات وتنظيمها:

فيما يخص البيانات النظرية (المكتبية) فـإن معالجتـها تتمثـل في الفحـص والتنظـيم والاستكمال ثم الصياغة والمراجعة والتوثيق، أما فيما يخص البيانات الميدانية، فإنه معالجتها تتضمن المراجعة المدققة للتأكد من أن تلك البيانات مستوفاة، وقد يترتب علـى ذلـك حذف بعض الحالات، واستكمال بيانات معينة أو إضافة حالات جديدة ...ويتم إدخال البيانات في الحاسوب باستخدام أحد البرامج الإحصائية المعروفة، وفي هذه المرحلة يتعين مراجعة البيانات التي تم إدخالها بهدف التأكد من أن تلك البيانات (المدخلة) هي بالضبط البيانات الصحيحة. وتتم معالجة البيانات إحصائياً وفق خطة تتناسب مع طبيعة المشكلة، وخصائصها الكمية والنوعية وهدف البحث، علماً بأن التحليل الذي يـصلح لمعالجـة مشكلة ما قد لا يصلح لمعالجة مشكلة أخرى، ويتضمن التحليل الإحصائي خطوة أوليـة تتمثل في التوصل إلى وصف دقيق للبيانات قد يتمثل الوصـف الإحـصائي في معرفـة التكرارات والنسب (Percentages) للبيانات ومتوسطاتها المختلفة أو نزعتـها المركزيـة وتلخيصها في صورة موجزة توضح أهم خواصها، بحيث يصل الباحث إلى وصف شامل للمتغيرات وتنظيمها في جداول ورسوم بيانية، ومنحنيات وأشكال توضيحية تبين معالم المتغيرات وخصائص البيانات ، ويسمي هذا الميدان من ميادين علم الإحصاء بالإحـصاء الوصفي.

ومن خلال الوصف الإحصائي يتم التمهيد للتحليل الإحصائي المناسـب لأن الوصـف الإحصائي يوضح الخواص الإحصائية للمتغيرات، ومن ثم- بناءً على هذا الوصف- يتم تحديد نوع التحليل الإحصائي ومستواه، ومن الخطأ الاعتقاد أن المغالاة في اختيار الطرق الإحصائية المتناهية في دقتها يعني الوصول إلي نتائج قوية، ذلك أن نوع التحليل يعتمـد على طبيعة البيانات وخصائصها والموضوع الذي يتم بحثه، فبعض الموضوعات لا تحتاج سوى إلى تحليل بسيط لأنها بطبيعتها لا تحتمل تحليلات إحصائية متقدمة، وقد تكـون البيانات الوصفية غير حساسة للفروق المتناهية في الدقة (فلا يعقل مثلاً حساب مـشاهدة التليفزيون بالثانية)

(٦) التفسير:

ينطوي تفسير النتائج على نوع من التعميم، ويجب ألا يجاوز هـذا التعمـيم حـدود البحث،لأنه يقوم على إطار تحدده عينة الأفراد الذين أجريت عليهم الدراسة والأدوات التي استخدمت فيها، والمعالجات الإحصائية التي تمت الاسـتعانة بهـا في التوصـل إلى النتائج...الخ، ومن الخطأ الشائع إجراء البحث في إطار معين محدد ثم تعمـيم النتـائج والتوصيات بما يفوق أو يتجاوز هذا الإطار، وعلى الباحث أن يلتزم حدود موضـوعه والنتائج التي توصل إليها

(٧) صياغة التقرير:

يبدأ التقرير من حيث بدأت المشكلة باختيارها وصياغتها، وينتهي إلي حيـث انتـهت بالتحليل الإحصائي والتفسير النهائي للنتائج. فالتقرير بهذا المعني يسجل خطوات البحث في تطورها خطوة تلو خطوة ليكون بذلك أقرب إلي الموضوعية العلمية والتنظيم المنطقي المتناسق. ويشترط في لغة البحث أن تكون واضحة موجزة موضوعية إلي الحـد الـذي تتخفف فيه من تأكيد الذات حتى لا تصطبغ بصبغة ذاتية تبعدها عـن الـروح العلمـي الصحيح، وغالبا ما ينتهي التقرير بملخص واضح عن المشكلة ونتيجة بحثها ومدى قوة أو ضعف هذه النتائج. فالتقرير بهذا المعنى يتضمن رؤية متكاملة لموضوع البحث وإجراءاتـه ونتائجه ومقترحاته ومصادره، وجوانب القوة والضعف في دراسة المشكلة، والموضوعات البحثية التي يستثيرها البحث والأفكار البحثية الجديدة التي كشف عنها البحث ، ومدى صلاحية هذه الموضوعات والأفكار للبحث، وبذلك يفتح التقرير النهائي آفاقاً جديـدة للبحث العلمي في مجال التخصص

ولا شك أن الاستخدام الواعي للإحصاء يتطلب العناية الشديدة بتصميم الدراسـة مـن كافة الجوانب بما في ذلك الفروض والعينة، وأدوات جمع البيانـات، والـوعي بطبيعـة المتغيرات المقاسة، فالبيانات التي لم يهتم الباحث بتحديد مستوي قياس متغيراتها وتعـرف شكل توزيعها، أو البيانات غير الدقيقة أو التي لا تتعلق بتساؤلات الدراسة أو فروضها لن تؤدي إلي نتائج ذات قيمة مهما كان أسلوب التحليل الإحصائي ومستواه، بمعنى أوضح،

فإن الأساليب الإحصائية في تحليل البيانات تستند إلي صيغ رياضية مجردة، والاستخدام المناسب لتلك الأساليب يتطلب فحصاً مدققاً للبيانات المطلوب تحليلها،مثلما يتطلب دقة الفروض وكفاءة تفعيلها في جمع البيانات وتحليلها وتفسيرها ، بحيث تكون الفروض مرشداً لجهود الباحث

ثالثاً: الإحصاء والقياس

القياس بمعناه العام تعبير رقمي عن الموضوع الذي يتم قياسه، والأرقام هنا أعداد نسميها درجات، بحيث تعبر عن كم معين (Quantity)، ويعتمد التعبير على النواحي الوصفية والنواحي الكمية، وتهدف النواحي الوصفية إلي الكشف عن وجود الصفة أو عدم وجودها، كالتعبير عن العلاقة بالتليفزيون بالمشاهدة أو عدم المشاهدة، أو التعبير عن المساحة في المواد المطبوعة بالسنتيمتر المربع ، وعن المساحة في الراديو والتليفزيون بالوقت أو المدة. وإذا كانت النواحي الوصفية تكشف عن وجود الصفة وتمايزها، فإن النواحي الكمية تهدف إلى الكشف عن درجة وجود الصفة، بمعنى القيمة الكمية لهذه الصفة، على سبيل المثال، فإن المقارنة الوصفية فيما يخص مشاهدة التليفزيون تتضمن تصنيف المفحوصين حسب المشاهدة أو عدم المشاهدة (يشاهد التليفزيون- لا يشاهد التليفزيون)، أما المقارنة الكمية، فتتضمن درجة المشاهدة، بمعنى تصنيف المفحوصين مثلاً إلى كثيفي المشاهدة Heavy Viewers ومتوسطي المشاهدة Medium Viewers وخفيفي المشاهدة Light Viewers وذلك حسب الوقت الذي يقضيه الفرد يومياً في مشاهدة التليفزيون. وهكذا تعتمد الطريقة الإحصائية على التصنيف الوصفي والكمي للظواهر المختلفة فهي بذلك تقسم الصفات إلي أنواع لها أهميتها بالنسبة لهدف البحث، ثم تقسمها إلي درجات تقاس بها كل صفة من تلك الصفات أي أنها تبدأ وصفية وتنتهي كمية

هذا التصنيف الوصفي والكمي من أهم دعائم المعرفة البشرية لأنه يلخص المعلومات المختلفة في قدر مناسب من المعلومات التي يمكن استيعابها، كما أن التصنيف يكشف عن العلاقات الجوهرية التي تربط الأشياء بعضها بالبعض الآخر، ويعتمد التصنيف على مدى تمايز الأشياء، وعلى تعميم هذا التمايز بحيث تنقسم الأشياء أو صفاتها إلي مجموعات بين كل مجموعة وأخرى فروق أساسية تبرز هذا الفصل القائم بينها، بحيث تضم كل مجموعة

الأفراد الذين يشتركون معاً في صفات أساسية تبرزهم جميعا معا في وحدة متآلفة، فالنوع الإنساني يشتمل على المميزات الرئيسية للجنس البشرى ويحول بين هذا الجنس والأجناس الأخرى بحيث لا تتداخل معه في هذا التقسيم.

والتمايز قد يكون حاداً فاصلاً، أو يكون متداخلاً بدرجة قليلة أو كبيرة، ومن أمثلة التمايز الحاد في الصفات: قراءة الصحف مقابل عدم قراءة الصحف، ومن أمثلة التمايز المتداخل بدرجة كبيرة الوقت الذي يقضيه الناس في مشاهدة التليفزيون، ولهذا يرصد هذا الوقت في سلسلة متصلة من الدرجات بحيث يمكن جمعها في فئات مثل من أقل من ساعة، من ساعة لأقل من ساعتين، ساعتين فأكثر

من جهة أخرى يجب أن يكون أساس التقسيم واضحاً وإلا تداخلت الأسس واختلط الأمر، فمن الخطأ تقسيم جمهور التليفزيون إلى ذكور وإناث وحضريين وإنما الصواب أن نقسم الجمهور حسب الجنس (ذكور -إناث)، ثم نعود لنقسم كل جنس حسب منطقة الإقامة (حضر– ريف)، حتى نستغرق الأقسام الفرعية، فالذكور قد يكونون حضريين أو ريفيين،كما أن الإناث قد يكن حضريات أو ريفيات. وهكذا نرى أن الأساس الأول للتقسيم في مثالنا هذا هو الجنس، والأساس الثاني للتقسيم هو منطقة الإقامة. والتقسيم قد يكون في صورة أقسام منفصلة أو متصلة، فالأقسام المنفصلة هي التي لا يمكن تجزئتها (فالفرد قد يكون ذكراً أو أنثى، وإما أن يكون مصرياً أو غير مصري...)، أما الأقسام المتصلة فهي التي يمكن تجزئتها (فالفرد قد يشاهد التليفزيون بدرجة كثيفة أو متوسطة أو ضعيفة) وذلك حسب المدة التي يقضيها يوميا في مشاهدة التليفزيون ، أي أن مدة المشاهدة تتوالي في تسلسل متصل من الارتفاع إلى الانخفاض... وهكذا تنقسم البيانات العددية بالنسبة لتمايزها إلي نوعين رئيسين: منفصلة ومتصلة.

أما حسب عدد التصنيفات، فإن التصنيف الإحصائي ينقسم إلي نوعين رئيسيين: الأول هو التصنيف الثنائي – بمعنى أن المتغير يقسم الصفة إلى تصنيفين فقط (ذكور مقابل إناث، ريف مقابل حضر....الخ)، أما التصنيف المتعدد فيعني أن المتغير يقسم الصفة إلى أكثر من تصنيف، فإذا افترضنا أن المتغير هو(مشاهدة التليفزيون)، فإن هذه المشاهدة فقد تكون: عالية، متوسطة، منخفضة، ويمكن تقسم مستويات المشاهدة حسب الجنس (ذكور وإناث)، كما يمكن إحداث تقسيمات أخرى في إطار كل تقسيم فرعي.

رابعاً:الإحصاء وتحليل البيانات

إن الكثير من البحوث العلمية ينتج عنها كميات كبيرة من البيانات التي يصعب تفسيرها في صورها الخام أو الأصلية، ونظراً لأن علم الإحصاء يتضمن الأساليب المتنوعة التي تستند إلى أسس رياضية أو منطقية وتصمم لمعاونة الباحث في تلخيص وتفسير البيانـــات الـــتي يحصل عليها، فقد انتشر استخدام الأساليب الإحصائية في مختلف المجالات البحثية .بما فيها مجالات العلوم الإنسانية، ولهذه الأساليب وظيفتان أساسيتان هما الوصف والاستدلال .

فالإحصاء الوصفي Descriptive Statistics يستخدم في تصنيف وتنظـــيم وتلخـــيص البيانات الكمية والنوعية بحيث يمكن فهمها وتفسيرها واستخلاص معلومات مفيدة منها، ويتعلق هذا النوع من الإحصاء بالحالات أو العينات التي بين يدي الباحث فالبيانـــات لا تكون ذات فائدة تذكر ما لم يتم تلخيصها في جداول وتمثيلها برسوم وأشـــكال بيانيـــة ووصفها باستخدام التكرارات والمتوسطات والتباين والارتباطات .

أما الإحصاء الاستدلالي Inferential Statistics فيستخدم في التوصل إلى استدلالات أو تعميمات على المجتمعات استناداً إلى بيانات متعلقة بعينات مستمدة من هذه المجتمعـــات وفقا لأساليب معاينات عشوائية .

وسواء كان الإحصاء وصفياً أو استدلالياً، فإن استخدامه في تحليل البيانات يتم في إطار تصميم عام يتعلق بالمتغيرات التي نهتم بقياسها، والمقارنات المطلوب إجراؤها....وذلـــك بهدف تيسير إمكانية التوصل إلى تفسيرات صادقة ومفيدة ومختصرة من بيانات كـــثيرة، ولكي يستخدم الإحصاء الوصفي أو الإحصاء الاستدلالي استخداماً مناســـباً في تحليـــل البيانات يجب أن ترتكز عملية جمع البيانات ومعالجتـــها وتنظيمهـــا وتفـــسيرها علـــى التساؤلات التي يريد الباحث الإجابة عنها أو الفروض التي يريد التحقق منها، مع الأخـــذ بالاعتبار طبيعة البيانات المستخدمة ومستويات المتغيرات المدروسة،،، إذ إن عدم مراعـــاة هذه الاعتبارات الأساسية هو الذي يؤدى إلى إساءة اســـتخدام الإحـــصاء في البحـــوث والتوصل إلى نتائج زائفة وتفسيرات مضللة

خامساً: استخدام الحاسوب في التحليل الإحصائي للبيانات:

كثيراً ما يتردد على السنة الباحثين في العلوم الاجتماعية تساؤل حول كيفية دراسة طرق تحليل البيانات والإحصاء دون وجود الخلفية الأساسية في الرياضيات التي تتصف بالرمزية والتجريد.

وهذا التساؤل بالطبع له ما يبرره، فمما لا شك فيه أن دراسة الرياضيات تيسر على الباحثين الفهم المستنير للأسس الرياضية التي تبنى عليها طرق وأساليب تحليل بيانات البحوث، غير أن ذلك لا يعني ضرورة أن يكون الباحث ماهراً في الرياضيات وفى استخدام أساليب المعالجات الرمزية حتى يستطيع إتقان الأساليب الإحصائية وطرق تحليل البيانات . إن استخدام الإحصاء في تحليل البيانات يحتاج إلى قدر من التفكير المنطقي في المشكلة البحثية وكيفية معالجتها إحصائياً، ويمكن للباحث أن يتقن ذلك سواء كان لديه خلفية قوية في الرياضيات أم لا ، فالمهم أن يكون لديه الرغبة في فهم الأساليب الإحصائية التي يمكن أن تساعده في تحليل بيانات بحثه للتوصل إلى نتائج سليمة، كما أن عملية تحليل البيانات تتطلب قدراً من العمليات الحسابية والجبرية ، وقد أدى التقدم الكبير الذي حدث في الآلات الحاسبة والحواسيب إلى تسهيل هذه الأمور

المشكلة أن الكثير من الباحثين لا يعطون الاهتمام الكافي لتحليل البيانات الميدانية بحجة أن الحاسوب سيجرى هذا التحليل عن طريق البرامج الإحصائية المتوافرة! لكن من الضروري هنا التأكيد على أن عملية تحليل البيانات تعد جزءا أساسياً لا يتجزأ من عملية البحث ذاتها، ويترتب عليها استخلاص النتائج والتفسيرات، ومما لا شك فيه أن الحاسوب يمكنه تخزين كميات كبيرة من البيانات يعجز العقل البشرى عن تخزينها والتعامل معها، كما يمكن للحاسوب إجراء العديد من العمليات الإحصائية والرياضية المعقدة على البيانات التي ربما يحتاج العقل البشرى إلى وقت طويل وجهد شاق لإنجازها، فالحاسوب يتميز بالسرعة الفائقة والدقة المتناهية في إجراء هذه العمليات مما يوفر على الباحث كثيراً من الوقت والجهد. ومن المعروف أن الباحث في العلوم الاجتماعية بعامة لا يمكنه دراسة الظواهر السلوكية دراسة متعمقة دون أن يأخذ بالاعتبار المتغيرات الأساسية التي يمكن أن تؤثر في هذه الظواهر، وهذا يجعل البيانات التي يجمعها الباحث عن ظاهرة معينة متعددة

المتغيرات، وهذا يتطلب الإفادة من إمكانات الحاسوب في تحليل هذه البيانات وتمثيلـــها بمختلف الصيغ حتى يتسنى له دراستها دراسة فاحصة وتفسيرها تفسيراً علمياً .

وعلى الرغم من ذلك فإن الحاسوب لا يستطيع التمييـــز بـــين الاســـتخدام المناســـب والاستخدام غير المناسب للأساليب الإحصائية نظراً لأنه يجرى العمليات التي تتطلبها هذه الأساليب بطريقة آلية وفق نظام معين على أي بيانات يدخلها الباحث، فمثلاً يمكـــن أن يوجد البرنامج متوسط متغير اسمي مثل النوع أو الديانة أو الانتماء إلى نادى معين ما دام الباحث قام بترميز أي من هذه المتغيرات ترميزا كمياً ، غير أن النتائج في هـــذه الحالـــة تكون عديمة المعنى وذلك لأن المدخلات غير المناسبة للأسلوب الإحصائي تؤدى عادة إلى مخرجات لا تعني شيئاً ، وهو ما يعرف لدى المشتغلين بعلـــوم الحاســـوب بـــمـــصطلح Garbage-in, Garbage-out وهذا يقتضي ألا يحاول الباحـــث اســـتخدام أســـلوب إحصائي معين دون فهم متطلبات هذا الأسلوب وفرضياته وملاءمتـــه لنـــوع البيانـــات وطبيعتها ومعنى المعطيات الإحصائية التي تنتج عن استخدامه .

وهناك العديد من البرامج والحزم الإحصائية المتوافرة التي يمكن أن يستخدمها الباحـــث في تحليل البيانات، ومن أهم البرامج شائعة الاستخدام في العلوم الاجتماعية والإنسانية نخص بالذكر:

(١) حزمة البرامج الإحصائية للعلوم الاجتماعيةSPSS :

تعد الحزمة الإحصائية للعلـــوم الاجتماعيـــة Statistical Package for the Social Sciences من أكثر الحزم استخداماً في تحليل البيانات حيث تشتمل على معظم الأساليب الإحصائية التي تستخدم في العلوم الاجتماعية، ويمكن إدارة هذا البرنامج في بيئة النوافـــذ Windows بمرونة كبيرة حيث يتم التفاعل بين المستخدم والبرنامج أثناء عمليات التحليل Interactive Computing مما يتيح للباحث تأمل بياناته وإعادة فحصها وتحليلها بطـــرق متعددة .

(٢) حزمة التحليل الإحصائي MINITAB

وهى من الحزم الحاسوبية الجيدة التي يمكن إدارتها في بيئة النوافذ ويمكن باستخدامها تحليل البيانات بمختلف أنواعها بمرونة وسرعة وتمثيل البيانات بيانياً والتخاطــب مــع الأوامــر وتحريك البيانات لدمجها أو فصلها مما يمكن الباحث من اســتخلاص أكــبر قــدر مــن المعلومات وهذه الحزمة من تصميم مؤسسة MINITAB الأمريكية .

(٣) نظام تحليل البيانات Statistical Analysis System (SAS)

ويعد من الحزم الإحصائية البسيطة والقوية التي تفيد الباحثين في تحليل البيانات في المجالات التطبيقية المختلفة، ويتميز هذا النظام بالمرونة وسهولة الاستخدام، ويجرى العديــد مــن الأساليب الإحصائية الوصفية والاستدلالية المتقدمة التي يحتاجها الباحثون في علوم النفس والتربية والاجتماع وإدارة الأعمال والطب وغيرها ويستخرج النتائج بدقة، وهذا النظــام من تصميم مؤسسة SAS بولاية شمال كارولينا الأمريكية .

(٤) برنامج تحليل بيانات المسوح MICROTEST

وهو من البرامج التي تفيد في تحليل البيانات المستمدة من مقاييس الاتجاهــات والميــول، والمسوح التربوية والاجتماعية والاقتصادية وبنظم البرنامج البيانات في جداول مستعرضة أو مصفوفات ويمثلها بيانياً ويجرى عليها العديد من التحلــيلات الإحــصائية الوصــفية والاستدلالية . والبرنامج من تصميم مؤسسة نظــم الحاســوب National Computer Systems الأمريكية .

سادساً:المتغيرات والمفاهيم من المنظور الإحصائي:

(أ) المتغيرات:

المتغير (Variable) بالمعنى البحثي والإحصائي هو الخاصية أو الصفة عند مجموعة أو عينة الأفراد أو الظاهرة الخاضعة للبحث، وهناك أسس متعددة لتصنيف المتغيرات، غير أن مــا يعنينا في هذه الجزئية- ليس أسس التصنيف- ولكن المتغيرات ذاتها:

المتغيرات الكمية والمتغيرات النوعية:

المتغيرات الكمية Quantitative هي تلك المتغيرات التي تعبر عـــن مقـــدار (Quantity) بحيث يمكن ترتيب الأفراد من الأصغر إلى الأكبر، أو من الأقل إلى الأكثر، مثال ذلـــك: غدد ساعات مشاهدة التليفزيون، مستوى المعرفة، مستوى التعليم.... الخ. هذه المتغيرات وما شاهها هى متغيرات كمية، وهى قد تكون متصلة أو منفصلة، فـــالمتغيرات الكميـــة المتصلة هي التي تأخذ قيمة موجبة أو سالبة، صحيحة أو كسرية مثل (الوقت المنقـــضي يوميا في استخدام الإنترنت، أو في مشاهدة التليفزيون..... أما المتغيرات المنفصلة فهـــي تأخذ قيماً صحيحة فقط مثل عدد المجلات أو عدد الصحف التي يطالعها الفرد. في مقابل ذلك فإن المتغيرات النوعية (Qualitative) إنما تعبر عن خاصية معينة من حيث وجودها أو عدم وجودها كالجنس (ذكر، أنثى)، ومحل الإقامة (ريف، حضر)، والتخصص العلمي (صحافة، إذاعة، كيمياء، رياضيات ... الخ) أي أنه لا يمكن ترتيب الأفراد من الأصـــغر إلى الأكبر حسب هذه الخاصية.

المتغيرات المستقلة والمتغيرات التابعة:

هناك فكرة أساسية تقوم على أن المتغير هو كل ما يؤثر أو يحتمل أن يؤثر في الظـــاهرة المدروسة، فالجنس Gender قد يؤثر في التخصص الدراسي، بمعنى أن الإناث قـــد يكـــنَّ أكثر إقبالاً على تخصصات معينة مقارنة بالذكور، أو العكس، والمستوى التعليمي للأبوين قد يؤثر في الدور الوالدي ... وهنا يكون الجنس، والمستوى التعليمي متغيرات مـــستقلة (independent) تمييزاً لها عن المتغيرات التابعة (dependent) فالمتغيرات المستقلة هي التي يتعامل معها الباحث إما كحقائق معطاة، أو معروفة وإما باعتبارها مؤثرات محتملـــة في الظاهرة أو في المتغير التابع، فالباحث قد يتعرف على المستوى التعليمي للوالدين، أي ما إذا كان جامعياً مثلاً، أو دون الجامعي ... الخ، ويتعامل معه كمتغير مستقل يمكن أن يؤثر في الظاهرة مجال الدراسة، وبموجب بعض الطرق الإحصائية يمكن للباحث معرفة حجـــم وطبيعة هذا التأثير، كما يمكنه عزل أو تحييد هذا التأثير، أو تفاعل تأثير متغير المـــستوى التعليمي للوالدين مع متغير آخر أو متغيرات أخرى. أما المتغيرات التابعة فهي المتغيرات التي

يحددها الباحث باعتبارها متغيرات واقع عليها التأثير... فعند دراسة العلاقة بين المستوى التعليمي للأب وأساليب التنشئة الأسرية تكون هذه الأساليب بمثابة المتغيرات التابعة، وهى التي يحددها الباحث ويصيغها في صورة بنواد أسئلة أو أبعاد، أما المستوى التعليمي فهـــو المتغير المستقل الذي يرصده الباحث.

المتغيرات من منظور القياس الإحصائي :

سبقت الإشارة إلى هذه الفكرة، لكننا سنشير إليها بشكل أوضح. فمن مـــن منظـــور القياس الإحصائي هناك المتغيرات الاسمية، والمتغيرات الرتبية، والمتغيرات الفترية، والمتغيرات النسبية:

المتغير الاسمي (Nominal):

إنه الخاصية أو النوعية بما يفيد التصنيف فقط (مثل: ذكور مقابل إنـــاث ، وكـــذلك : ريف مقابل حضر)، وعندما نستخدم الأرقام في التصنيف الإحصائي لتلك المتغيرات، فإن الأرقام يقتصر دورها على مجرد التصنيف دون أي دلالة كمية، على سبيل المثال قد نرمز إلى الذكور بالرقم (١) ونرمز إلى الإناث بالرقم (٢)، فالرقم (٢) لم يستخدم بمعنى أنه أكبر من الرقم (١)، كما أن الرقم (١) لم يستخدم بمعنى أنه أصغر من الرقم (٢) — وإنما استخدم كلاً منهما لتصنيف الأفراد إلى ذكور وإناث.

المتغير الرتبي Ordinal:

وهو تصنيف الأفراد أو الموضوعات من خلال ترتيبهم تصاعدياً أو تنازلياً، مثال ذلـــك ترتيب أفراد البحث تنازلياً حسب الفئة العمرية، بصرف النظر عن الفروق بين مستويات التصنيف، فالشخص الأكبر عمراً يأخذ ترتيباً أكبر من الشخص الأوسط عمراً، وهـــذا بدوره يأخذ ترتيباً أكبر من الشخص الأصغر عمراً، والشخص الذي يليه في العمر يأخـــذ ترتيباً أقل ....وهكذا. فإذا كانت الأرقام الترتيبية هي: ٥ ،٤،٣، ١،٢، فإن الـــرقم ٥ يرمز إلى الترتيب الأكبر، بينما الرقم ١ يرمز إلى الترتيب الأصغر، لكن ذلـــك لا يعـــني بالضرورة تساوي الفروق في الأعمار بين المفحوصين(فالفرق بين عمر الشخص الأول

وعمر الشخص الثاني قد يكون شهراً، أما الفرق بين الشخص الثاني والثالث فقد يكون عاماً....)

المتغير الفئوي أو الفتري (Interval):

ويتضمن معنى المتغير الرتبي (حيث يمكن ترتيب الأفراد من الأصغر إلى الأكـبر، أو مـن الأكبر إلى الأصغر) ولكنه بجانب ذلك يسمح بتحديد الفروق بين هؤلاء الأفـراد في الصفات أو السلوكيات التي يتم قياسها، والتعرف الدقيق على المسافة بين هؤلاء الأفراد، وما إذا كانت متساوية أو غير متساوية، أي أن لهذا المتغير وحدة قياس، إلا أن الصفر أو نقطة الأصل أو البداية لا تعنى غياب الظاهرة أو الخاصية المقاسة، فحصول بعض الطلاب على (صفر) في مقرر معين لا يعنى أنهم لا يعرفون شيئاً في هذا المقرر، وإنما هم (بالتأكيد) لديهم بعض المعلومات عنه، فالصفر هنا صفر نسبى أي غير مطلق. وعند هذا المستوى من المتغيرات – وأقصد بها المتغيرات الفئوية أو الفاصلة، تعمل البحوث في مجـالات العلـوم الاجتماعية والإنسانية عموماً، فهذه العلوم لم تعتمد بعد على المتغير النسبي (Ratio) على النحو المعمول به في العلوم الطبيعية.

المتغير النسبي (Ratio):

وهو المتغير الذي تتوفر فيه معاني المتغير الفتري بجانب كونه يتضمن معنى الصفر المطلق، بمعنى غياب الصفة المقاسة، والمتغيرات من هذا النوع تقع ضمن العلوم الطبيعية، ولـيس العلوم الاجتماعية،فنحن نقول إن درجة الحرارة تساوي الصفر، في حين لا يمكن القـول إن مشاهدة الشباب للتليفزيون تساوي الصفر (حتى ولو أفـادوا بـأنهم لا يـشاهدون التليفزيون)

(ب) المفاهيم:

المفاهيم ببساطة هي تجريدات (Abstracts)، إنها بناءات صغيرة جزئية، وهـي اللبنـات الأولية التي تساعد في بناء النظريات والقوانين العلمية التي هي هدف العلم، وبدون تجريد المفاهيم عن الظاهرة لا يمكن إيجاد بناء نظري. فالتجريد إذن بمثابة فصل وتحديد الظواهر أو المتغيرات والتعبير عنها بواسطة (مفاهيم) معينة تناسب مجال التخصص العلمي وتميـزه.

وتساعد المفاهيم على نقل الاستنتاجات والمشاهدات والنتائج بين المتخصصين في المجـــال الواحد، وكذلك بينهم وبين القراء من خارج هذا المجال.

ومن الناحية اللغوية والدلالية فإن (المفهوم) عبارة عن تركيب لغوى منطقي أو مـــسمى معين يدل على شيء أو ظاهرة أو معنى، ويتم التوصل إلى المفهوم من خلال الانطباع أو الملاحظة أو التجربة ... الخ، ومن خصائص المفهوم أنه يتضمن تركيزاً شديداً لمعلومـــات وجوانب عديدة. إن مفهوم (الاتصال) على سبيل المثال يتكون من كلمة واحدة تتركز فيها معان عديدة تشمل أشكال الاتصال ومستوياته وتأثيره وعملياته....الخ.

والمفاهيم – حتى في الحياة العادية – تمكننا من إضفاء المعنى على العالم المحيط بنا، فمـــن خلال المفاهيم يصبح للحقيقة معنى محدد، والمفاهيم هي الوسائل التي من خلالها نصل إلى خبراتنا، أي أن إدراكنا للعالم يعتمد اعتماداً كبيراً على رصيدنا من المفاهيم التي نمتلكها، ويمكننا التحكم فيها، فكلما زادت مفاهيمنا، زادت قدرتنا على تجميع بيانات ذات معنى، وكذلك على التأكد مما ندركه وما نعرفه عن هذا الكون، وهذا يعنى أن النظرة إلى الحقيقة الموضوعية الواحدة تختلف من فرد إلى آخر تبعاً لما لديه من مفاهيم، إن الطبيب عنـــدما يشخص مرضاً معيناً فإنه يبنى تشخيصه على مجموعة مفاهيم تختلف اختلافاً كبيراً عمـــا لدى الرجل العادي من مفاهيم عن هذا المرض.

ولقد أضفي العلماء في تخصصاتهم معان محددة على مجموعة من المفاهيم، وهذه المفـــاهيم تشكل إدراك الظواهر بطريقة خاصة بحيث يستطيع الباحث في مجال معـــين أن يفهـــم الجزئية التي يريد دراستها، وتتعدد المفاهيم المستخدمة في بحوث الاتصال مثل: التعرض، الانتقائية، الوسيلة، الرسالة، الجمهور، التأثير....الخ

وتدل الكثير من المفاهيم على معان مختلفة، فالمفهوم الواحد قد يعنى شيئاً معيناً في أحـــد العلوم، ولكنه يعنى شيئاً مختلفاً في علم آخر، فكلمة (المسرح) في مجال الأدب والفن، تعنى الفن المسرحي المعروف، أما في العلوم العسكرية فإن كلمة (مسرح) كثيراً ما تدل مكان القتال أو المعركة (مسرح العمليات)، وكلمة (جاذبية) في علم الفيزياء تعـــنى الجاذبيـــة الأرضية أو المغناطيسية، أما في علم النفس فقد تدل على الإعجاب أو الحب بين الناس.

كما أن المفهوم الواحد قد يدل على أكثر من معنى في المجال الواحد، فمفهوم (وظيفة) قد يدل على المهنة أو العمل الذي يزاوله الشخص، كما قد يدل على عمليات نفسية داخلية أي داخل الفرد، كالتفكير والتأمل. في الوقت نفسه، هناك مفاهيم متقاربة في المعنى مثل: (الحافز)، (الدافع)، (الحاجة) إن هذه المفاهيم الثلاثة متقاربة المعنى في علم النفس، وكذلك مفهوم (الرأي) ومفهوم (الاتجاه). لكل ذلك، تبرز ضرورة تعريف (المفاهيم) المستخدمة في البحث العلمي، وإلا حدثت الفوضى وظهر سوء الفهم. من هنا تأتي ضرورة تعريف المفاهيم التي وردت في فروض البحث أو تساؤلاته.

وهناك نوعان من التعريفات، الأول: نظري، والثاني إجرائي. إن التعريف النظري يتمثل في تعريف المفهوم كبنية فكرية حسبما تتضمن ذلك الكتابات التي تناولت هذا المفهوم، فالباحث الذي يقوم بدراسة الرأي العام من زاوية معينة، سوف يجد تعريفات متعددة لمفهوم (الرأي العام) في الكتابات النظرية ومن ذلك يستمد الباحث تعريفاً نظرياً لمفهوم الرأي العام.

أما التعريف الإجرائي، فإنه التعريف الذي يضعه الباحث للتعبير عن المفهوم. فالتعريف الإجرائي يعبر عن الرؤية الخاصة بالباحث بشأن كيفية تعامله مع المفهوم وكيفية قياسه، فقد يضع الباحث تعريفاً للرأي العام بأنه الرأي الذي عبرت عنه غالبية عينة البحث تجاه سيطرة الدولة على وسائل الإعلام. والتعريف الإجرائي للمفهوم يدل القارئ على ما يقصده الباحث، وبالتالي يكون المعنى واحداً للقارئ والباحث على السواء، كما أن هذا التعريف يمثل همزة الوصل بين النظرية والواقع، إنه يجعل المفاهيم قابلة للقياس والاختبار، فالتعريف الإجرائي للذكاء مثلاً يقود إلى كيفية قياسه، والتعامل معه تجريبياً.

وتتضح أهمية التعريفات الإجرائية في البحث العلمي أيضاً من واقع ارتباط البحوث بمجتمع معين، وواقع معين بخصائص معينة قد لا تناسبه التعريفات النظرية للمفاهيم، ومن ثم يستفيد الباحث من هذه التعريفات النظرية في بلورة تعريف إجرائي بلائم الواقع الذي يجرى فيه البحث. ويستنتج من ذلك أن التعريف الإجرائي يجب ألا يتناقض مع الأصول

النظرية المتعارف عليها في التخصص، بل إن هذه الأصول يجب أن يكون بمثابة مـــصدر الإمداد الفكري للحصول على صياغة تعريف إجرائي يناسب البحث.

الفصل الثاني
معامل الارتباط

المبحث الأول

معنى الارتباط وخصائصه

الارتباط (Correlation) بمعناه الإحصائي هو اقتران التغير في ظاهرة معينة بـــالتغير في ظاهرة أخرى، بمعنى أنه إذا تغيرت الظاهرة (س) تتغير الظاهرة (ص)، وقد يكون الـــتغير موجباً أو سالباً. فالتغير الموجب يعني زيادة في الظاهرة (س) عندما تزداد الظاهرة (ص)، أو زيادة في الظاهرة (ص) عندما تزداد الظاهرة (س) . أما التغير السالب فيعني نقـــصان في الظاهرة (س) عندما تنقص الظاهرة (ص)، أو نقصان في الظاهرة (ص) عنـــدما تنقص الظاهرة (س)، وقد يكون التغير عكسيا، بمعنى زيادة الظاهرة (س) عندما تتنـــاقص الظاهرة (ص) ، أو تناقص الظاهرة (س) عندما تزيد الظاهرة (ص) .....هـــذا الـــتغير الاقتراني يقاس بمعاملات الارتباط والتي لها أهمية قصوى في التحليل الإحصائي بمستوياته المختلفة. وتتلخص أهم الخواص الإحصائية لمعاملات الارتباط مـــن خـــلال النقـــاط الأساسية التالية:

١-حدوث الارتباط

يصل الارتباط إلى نهايته العظمي عندما يقترن تغير المتغير الأول اقتراناً تاماً بالتغير في المتغير الثاني، هذا الارتباط التام قد يكون موجباً أو سالباً، كما سبقت الإشارة ، ومن أمثلـــة الارتباط التام الموجب اقتران زيادة وقت مشاهدة التليفزيون بزيادة وقت الفراغ بحيـــث يظل ترتيب الأفراد بالنسبة للمتغيرين ثابتاً، فكلما زاد وقت مشاهدة التليفزيون زاد وقت الفراغ، وكلما زاد وقت الفراغ زاد وقت مشاهدة التليفزيون، ويظل ترتيب الأفراد كما هو حسب كلا المتغيرين، وتكون قيمة الارتباط التام الموجب مساوية للواحد الـــصحيح (+١)، أما الارتباط التام السالب فمن أمثلته اقتران زيادة الوقت المنقـــضي في مـــشاهدة التليفزيون بنقصان الوقت المنقضي في التحدث مع أفراد الأسرة بحيث يكـــون ترتيـــب الأفراد واحداً في المتغيرين (الوقت المنقضي في مشاهدة التليفزيون والوقـــت المنقـــضي في التحدث مع أفراد الأسرة)، وتكون قيمة الارتباط التام الـــسالب مـــساوية للواحـــد الصحيح (-١)

أي أن الحدود الحقيقية لقيمة معامل الارتباط تمتد من + ١ إلى – ١ ، أي من الارتباط الموجب التام إلى الارتباط السالب التام، وفيما بين ذلك قد يكون الارتباط طردياً أو موجباً (زيادة المتغير الأول بزيادة المتغير الثاني، أو زيادة المتغير الثاني بزيادة المتغير الأول)، كما قد يكون عكسياً أو سالباً (نقصان المتغير الأول بزيادة المتغير الثاني، أو زيادة المتغير الثاني بنقصان المتغير الأول)، وقد يكون الارتباط غير خطي (علاقة منحنية) بين المتغيرين، وقد تصل قيمة معامل الارتباط إلى الصفر، وفي هذه الحالة لا يكون هناك ارتباط بين المتغيرين. نخلص من ذلك إلى أن معامل الارتباط بين متغيرين أو ظاهرتين يتخذ أحد الِأشكال الستة (ارتباط طردي تام، ارتباط عكسي تام ، ارتباط طردي، ارتباط عكسي، ارتباط غير خطي، عدم وجود ارتباط)، فإذا طبقنا هذه الفكرة على العلاقة بين الوقت اليومي المنقضي في قراءة الصحف (س)، والوقت اليومي المنقضي في مشاهدة التليفزيون (ص)، فإن معامل الارتباط بين هذين المتغيرين قد يكون:

- موجباً تاماً: أي أن قيمة معامل الارتباط تساوي +١ وهذا يعني أن الوقت اليومي المنقضي في قراءة الصحف (س) يزيد بزيادة الوقت اليومي المنقضي في مشاهدة التليفزيون (ص) وذلك بمقدار ثابت أو متساوي (فإذا زاد أحد المتغيرين بمقدار وحدة واحدة، زاد المتغير الآخر بمقدار وحدة واحدة، وإذا زاد أحد المتغيرين بمقدار وحدتين، زاد المتغير الآخر بمقدار وحدتين.... وهكذا)

- موجباً: أي أن قيمة معامل الارتباط تكون موجبة وإن كانت تقل +١ وهذا يعني أن الوقت اليومي المنقضي في قراءة الصحف (س) يزيد بزيادة الوقت اليومي المنقضي في مشاهدة التليفزيون (ص) ولكن بمقدار غير ثابت أو غير متساوي

- سالباً تاماً: أي أن قيمة معامل الارتباط تساوي -١ وهذا يعني أن الوقت اليومي المنقضي في قراءة الصحف (س) ينخفض بزيادة الوقت اليومي المنقضي في مشاهدة التليفزيون (ص) وذلك بمقدار ثابت أو متساوي (فإذا نقص أحد المتغيرين بمقدار وحدة واحدة، زاد المتغير الآخر بمقدار وحدة واحدة، وإذا نقص أحد المتغيرين بمقدار وحدتين، زاد المتغير الآخر بمقدار وحدتين.... وهكذا)

- سالباً : بمعنى أن قيمة معامل الارتباط تكون سالبة وتقل عن -١ وهـذا يعـني أن الوقت اليومي المنقضي في قراءة الصحف (س) ينخفض بزيادة الوقت اليومي المنقـضي في مشاهدة التليفزيون (ص) لكن قيمة الانخفاض والزيادة غير ثابتتين أو متساويتين)
- ارتباط غير خطي، أو علاقة منحنية، بمعنى أن الوقت اليومي المنقضي في قراءة الصحف (س) ينخفض بزيادة الوقت اليومي المنقضي في مشاهدة التليفزيـون (ص) وذلـك بمقدار غير ثابت.

ب- زيادة أو نقصان الدرجات بكمية ثابتة

لا يتأثر معامل الارتباط بزيادة أو نقصان درجات الاختبار بكمية ثابتة، فإذا أضفنا عدداً ثابتاً إلى جميع درجات أي اختبار فإن هذه الإضافة لا تؤثر في ترتيب الأفـراد بالنـسبة لدرجات الاختبار ويبقي التغير الاقتراني القائم بين الظاهرتين كما هو. المنطق نفسه إذا طرحنا عدداً ثابتاً من جميع درجات أي اختبار فإن هذا النقصان لا يـؤثر في الترتيـب، ويمكن أن نستعين بهذه الفكرة في تبسيط العمليات الحسابية وذلك بإضافة أو بطرح عدد ثابت من درجات الاختبارات التي تحسب معاملات ارتباطها، أي أن معامل الارتبـاط لم يتغير بطرح أو إضافة رقم ثابت

جـ- متوسطات معاملات الارتباط

يميل التوزيع التكراري لمعاملات الارتباط إلى الالتواء، وخاصة عندما تزداد القيم العدديـة لتلك المعاملات، لكن التوزيع التكراري لمعاملات الارتباط يقترب من التوزيع الاعتدالي كلما اقتربت القيم العددية للارتباط من الصفر، ويلتوي التوزيع التكراري التواء شـديداً كلما اقتربت الارتباطات من الواحد الصحيح. وقد لجأ فيشر R A. Fisher إلى تحويـل القيم العددية لتلك المعاملات إلى صورة جديدة تقيم عوج (التواء) ذلك التوزيع وتنحو به إلى التوزيع الاعتدالي، وتتلخص طريقة فيبشر في تحويل معاملات الارتباط إلى معـاملات لوغاريتمية تعتدل في توزيعها التكراري، وعندما تقل قيمة معامل الارتباط عـن ٠,٣٥ فإنها تصبح مساوية لقيمة المعامل اللوغاريتمي للارتباط ولذلك لا تحسب تلـك القـيم اللوغارتمية إلا إذا زادت قيمة معامل الارتباط على ٠,٣٥ ولهذه الفكرة الإحصائية أهمية

كبيرة في حساب متوسطات معاملات الارتباط وذلك لأن الالتــواء الــشديد للتوزيــع التكراري يؤثر على دقة متوسط التوزيع، ولذا تحول معاملات الارتباط (ر) إلى مقابلاتها اللوغارتمية (ز)، ثم يحسب متوسط القيم العددية للمقابلات اللوغارتمية ثم يحــول هــذا المتوسط إلى صورته الأصلية وهي معامل الارتباط (المقابل الارتباطي).

وبما أن عملية تحويل (ر) إلى (ز) تستغرق وقتاً وجهداً كــبيرين، لــذلك رصــدت المقابلات اللوغارتمية(ز) للارتباط (ر) في جداول إحصائية خاصة، والمثال بالجــدول الآتي الذي يوضح طريقة حساب متوسط معاملات الارتباط بطريقة المقابلات اللوغارتمية ومقارنة نتائج هذه الطريقة بقيمة الارتباط التي تقابل هذا المتوسط :

معاملات الارتباط (ر)	المقابلات اللوغارتمية (ز)
٠,٧٥	٠,٩٧
٠,٧٨	١,٠٥
٠,٨٣	١,١٩
٠,٩٤	١,٧٤
٠,٩٥	١,٨٣
مج ر = ٤,٢٥	مج ز = ٦,٧٨ م ز = ١,٣٥٦
م ر = ٠,٨٥	م ز = ٠,٨٨

يدل العمود الثاني من هذا الجدول على المقابلات اللوغارتمية لكل معامل من معــاملات ارتباط بالعمود الأول، على سبيل المثال إذا كانت قيمة معامل الارتباط (ر) تــساوي

٠,٧٥ فإن المقابل اللوغاريتمي (ز) لهذه القيمة هو ٠,٩٧ وهكذا بالنسبة لبقية معاملات الارتباط بالجدول ، وقد حسب متوسط معاملات العمود الأول فظهر أنه يساوي ٠,٨٥، أما متوسط المقابلات اللوغارتمية فيساوي ١,٣٥٦ هذا المتوسط يقابله معامل ارتبـــاط قدره ٠,٨٨ أما متوسط معاملات الارتباط في العمود الأول فهو ٠,٨٥ وهكذا ندرك أن الفرق بين المتوسطين يساوي ٠,٨٨-٠,٨٥=٠,٣

المبحث الثاني

أهم معاملات الارتباط

إن اختيار معامل الارتباط المناسب يتطلب تحديد نوعية المتغيرات المطلـــوب حـــساب الارتباط بينها، وقد سبقت الإشارة إلى أن هناك أساليب متعددة لقياس الارتباط، ويتحدد استخدام معامل الارتباط في ضوء طبيعة المتغيرات المطلوب إيجاد الارتباط بينها، أي ما إذا كانت تلك المتغيرات اسمية Nominal أو رتبية Ordinal أو فاصلة Interval أو نسبية Ratio وإذا كانت المتغيرات النسبية Ratio تناسب العلوم الطبيعية ، فإن المتغيرات الثلاثة الأخرى هي التي يشيع استخدامها في العلوم الإنسانية، وإذا كان الارتباط يعبر عن التغير الاقتراني بين متغيرين Two variables فإن هذا يعني أن الارتباط قد يكون بين: متغيرين اسميين، أو متغيرين رتبيين، أو متغيرين فتريين، كما قد يكون معامل الارتباط بين متغير اسمي والآخر رتبي، أو بين متغير اسمي والآخر فتري، أو بين متغير رتـــبي والآخـــر فتري. في هذه الجزئية سنناقش أهم معاملات الارتباط الشائعة بحيث يمكـــن للبـــاحثين اختيار الأساليب التي تناسب المتغيرات التي تتضمنها أبحاثهم، وهذا يعني من الاستفادة من هذه الجزئية يتطلب معرفة المقصود بالمتغيرات وخصائصها. وفيما يلي عرض موجز لأهم أساليب الارتباط من خلال أربعة عناوين أساسية هي:

- الارتباط بين متغيرين اسميين
- الارتباط بين متغيرين رتبيين
- الارتباط بين متغيرين أحدهما رتبي والآخر اسمي
- الارتباط بين متغيرين من النوع الفتري (معامل ارتباط بيرسون)

أولاً: الارتباط بين متغيرين اسميين

تتيح البرامج الإحصائية مجموعة أساليب لقياس الارتباط بين متغيرين اسميين (Nominal) ، وتتمثل تلك الأساليب في: معامل الاقتران الرباعي ، معامل ارتباط فاي، معامل التوافق، معامل كرامير ، معامل تشيبرو ، اختبار كا^2 للاستقلالية ، معامل لامدا ، وفيما يلي تعريف موجز بتلك الأساليب

١-معامل الاقتران الرباعي

يستخدم معامل الاقتران الرباعي Tetra choric Coefficient of Association **في حالة** المتغيرات النوعية (الاسمية) التي يقسم كل منها إلى وجهين فقط مثل (يشاهد – لا يشاهد) ، (يمتلك– لا يمتلك)، (يقرأ – لا يقرأ)، (يستمع – لا يستمع)....الخ، فإذا كان لدينا المتغيران النوعيان : س ، ص وكان المتغير (س) ينقسم إلى قسمين، وكذلك المتغير (ص) ينقسم إلى قسمين، فإنه يمكن الحصول على جدول توافق Contingency Table في صورة رصد مزدوج Cross – Tabulation

ص	س	
	قسم أول س	قسم ثان س
قسم أول ص	أ	ب
قسم ثان ص	حـــ	د

حيث أ، ب، حـــ، د ترمز إلى المشاهدات في صورة تكرارات موزعة على الأقسام المختلفة للمتغيرين (س ، ص) ، وفي هذه الحالة يتم حساب العلاقة بين المتغيرين س، ص باستخدام القانون:

$$\text{قت} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\text{أ} \times \text{د} + \text{ب} \times \text{ج}}$$

حيث قت : معامل الاقتران

أ: عدد الأفراد الذين يقعون في القسم الأول ص والقسم الأول س

ب: عدد الأفراد الذين يقعون في القسم الأول ص والقسم الثاني س

ج: عدد الأفراد الذين يقعون في القسم الثاني ص والقسم الأول س

د: عدد الأفراد الذين يقعون في القسم الثاني ص والقسم الثاني س

مثال:

الجدول الآتي يوضح توزيع عينة قوامها ٥٢ شخصاً من القائمين بالاتصال في إدارات الأخبار بالتليفزيون حسب الرأي في استخدام تقنية حديثة في المونتاج وزيادة جودة البرامج :

استخدام التقنية الحديثة	جودة البرامج	
	تحققت	لم تتحقق
نعم	٢٤	١٠
لا	٣	١٥

فالخلية الأولى (أ) تتضمن ٢٤ تكراراً، أما الخلية الثانية (ب) فتتضمن ١٠ تكرارات بينما تتضمن الخلية الثالثة (ج) ٣ تكرارات ، أخيراً فإن الخلية (د) تتضمن ١٥ تكرارا ولحساب معامل الاقتران الرباعي فإننا نطبق المعادلة السابقة فيكون:

$$\text{قت} = \frac{٢٤ \times ١٥ - ١٠ \times ٣}{٢٤ \times ١٥ + ١٠ \times ٣} = ٠,٨٥$$

ويتضح من ذلك ارتفاع قيمة الاقتران بين استخدام التقنية الحديثة وتحقق جودة البرامج، لكن القيم العددية لمعامل الاقتران الرباعي تميل إلى أن تكون أكبر من القيم العددية لمعاملات الارتباط الأخرى، ولذلك يتم تحويل قيمة معامل الاقتران الرباعي إلى قيمة معيارية (Z) وذلك باستخدام القانون:

$$Z = \frac{٢ \text{ قت}}{١ - \text{قت}^٢} \times \frac{١}{\sqrt{\frac{١}{\text{أ}} + \frac{١}{\text{ب}} + \frac{١}{\text{ج}} + \frac{١}{\text{د}}}}$$

وبتطبيق هذه المعادلة على المثال المذكور فإن قيمة Z تكون:

$$\frac{٢ \times ٠,٨٥}{١ - (٠,٨٥)^٢} \times \frac{١}{\sqrt{\frac{١}{٢٤} + \frac{١}{١٠} + \frac{١}{٣} + \frac{١}{١٥}}}$$

$$\frac{١,٧}{٠,٢٨} \times \frac{١}{\sqrt{٠,٥٤}}$$

$$= ٦,٠٧ \times ١,٤ = ٨,٥$$

وتقارن القيمة الناتجة– أي قيمة Z بالمقدار ± ١,٩٦ للتحقق من الدلالة عند مستوى ٠,٠٥ أو بالمقدار ± ٢,٥٨ للتحقق من الدلالة عند مستوى ٠,٠١ وذلك طبقاً لاختبار الذيلين (Two tails test) ، وتكون قيمة معامل الاقتران موجبة إذا كان أ × د > ب × جـ بينما تكون تلك القيمة سالبة إذا كان أ × د < ب × جـ . أما إذا كان أ × د = ب × جـ فإن قيمة معامل الاقتران تكون منعدمة أو تساوى صفراً

٢- معامل ارتباط فاي

يستخدم معامل ارتباط فاي phi coefficient إذا كانت البيانات تتعلق بمتغيرين ينقسم كل منهما انقساماً ثنائياً في صورة اسمية مثل (صواب – خطأ) ، (نعم – لا)، (ذكر – أنثى) (الرسوب – النجاح) ، ... أو حولت المتغيرات المتصلة إلى متغيرات ثنائية. فإذا أردنا تعرف العلاقة بين هذين المتغيرين فإن معامل ارتباط فاي phiيكون مناسباً، فكثيراً ما يكون لدينا إجابة ثنائية (نعم - لا) على أسئلة أو بنود في الاستبيان، مثل : هل تمتلك أكثر من جهاز تليفزيون... هل تقرأ جريدة معينة بانتظام..... هل شاركت في الانتخابات الأخيرة.....هل سافرت إلى الخارج....هذه الأسئلة وما شابهها من أسئلة ذات إجابة ثنائية (نعم & لا) يمكن إيجاد العلاقة بينها باستخدام معامل ارتباط فاي. قد يكون المتغير الأول هو مشاهدة التليفزيون يوميا (س) أما المتغير الثاني فهو قراءة الصحيفة بانتظام (ص)، و يمكن تصنيف الاستجابة على هذين السؤالين كما يلي:

المشاهدة	**القراءة (ص)**	
(س)	**نعم**	**لا**
نعم	أ	ب
لا	ج	د

حيث أ، ب، ج، د هي المشاهدات في صورة تكرارات تتوزع على الأقسام المختلفة لكل من مشاهدة التليفزيون يومياً (س) ، و قراءة الصحيفة بانتظام (ص)، ويمكن إيجاد العلاقة أو الارتباط بين استجابات المفحوصين على هذين المتغيرين باستخدام قانون معامل فاي:

$$\phi = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\sqrt{(\text{أ} + \text{ب})\ (\text{ج} + \text{د})\ (\text{أ} + \text{ج})\ (\text{ب} + \text{د})}}$$

حيث ϕ معامل فاي

أ = عدد الأفراد الذين أجابوا (نعم) على السؤال س وأجابوا (نعم) على السؤال ص

ب = عدد الأفراد الذين أجابوا (نعم) على السؤال س، وأجابوا (لا) على السؤال ص

ج = عدد الأفراد الذين أجابوا (لا) على السؤال س ، وأجابوا (نعم) على السؤال ص

د = عدد الأفراد الذين أجابوا (لا) على السؤال س وأجابوا (لا) على السؤال ص

مثال: الجدول الآتي يوضح الاستجابة على سؤالين، الأول يتعلق بقراءة إحدى الصحف اليومية بانتظام (س) ، والثاني يتعلق بقراءة إحدى المجلات الفنية الأسبوعية بانتظام (ص) كما هي موضحة بالجدول التالي:

ص \ س	نعم	لا
نعم	٨	٧
لا	١٥	١٠

والمطلوب حساب ارتباط استجابات أفراد العينة على هذين السؤالين ودلالته الإحصائية. **في هذا المثال نجد أن :**

قيمة (أ) = ٨

قيمة (ب) = ٧

قيمة (ج)= ١٥

قيمة (د) = ١٠

$$\text{ø} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\sqrt{(\text{أ} + \text{ب})\,(\text{ج} + \text{د})\,(\text{أ} + \text{ج})\,(\text{ب} + \text{د})}}$$

$$= \frac{\text{٨} \times \text{١٠} - \text{٧} \times \text{١٥}}{\sqrt{(\text{٨} + \text{٧})\,(\text{١٥} + \text{١٠})\,(\text{٨} + \text{١٥})\,(\text{٧} + \text{١٠})}}$$

$$= \frac{\text{٨٠} - \text{٣٥}}{\sqrt{(\text{١٥})\,(\text{٢٥})\,(\text{١٣})\,(\text{١٧})}}$$

$$= \frac{\text{٤٥}}{\sqrt{\text{١٤٦٦٢٥}}}$$

$$= \frac{\text{٤٥}}{\text{٣٨٢,٩}} = \text{٠,١١٧}$$

أي أن هناك ارتباطاً موجباً قدره ٠,١٢ تقريباً بين قراءة الصحيفة اليومية بانتظام وقراءة المجلة الفنية بانتظام ، أي أنه كلما زادت قراءة الصحيفة اليومية بانتظام زادت قراءة المجلة الفنية الأسبوعية بانتظام. ولمعرفة الدلالة الإحصائية لمعامل فاي عند مستوى معين أم لا علينا أن نحول قيمة فاي المحسوبة إلى Z بموجب القانون:

$$Z = ø \sqrt{\text{ن}}$$

فإذا كانت ø = ٠,١١٧ وكان إجمالي الأفراد (ن) = ٤٠

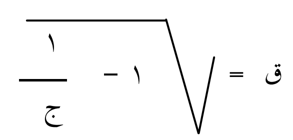
$$\text{فإن } Z = ٠,١١٧ \times \sqrt{٤٠}$$

$$= ٠,١١٧ \times ٦,٣ = ٠,٧٤$$

وعلينا أن نقارن قيمة Z الناتجة بالقيم اللازمة للدلالة الإحصائية ± ١,٩٦ أو ± ٢,٥٨ عند مستوى ٠,٠٥، ٠,٠١ على الترتيب وذلك بالنسبة لاختبار ذيلين.

٣- معامل التوافق:

وهو من الأساليب الإحصائية التي يمكن استخدامها للكشف عن العلاقة أو الارتباط بين متغيرات لا يمكن قياسها كمياً أو لا يمكن تصنيفها في رتب، وتكون هذه المتغيرات في صورة مجموعات أو أقسام يكون فيها لكل قسم صفة مميزة مثل (متزوج – أعزب – مطلق) لتمثيل الحالة الاجتماعية، وكذلك (منطقة حضرية- منطقة ريفية- منطقة صحراوية) لتمثيل منطقة الإقامة.... ولقياس العلاقة بين هذه المتغيرات يستخدم معامل التوافق Contingency Coefficient ، وهو أعم من معامل الاقتران الرباعي حيث يمكن تطبيقه سواء كان الجدول المزدوج من النمط ٢×٢ أو أكثر ، وسواء تساوى عدد خلايا الصفوف وعدد خلايا الأعمدة أو لم يتساوى، والقانون المستخدم للحصول على قيمة معامل التوافق هو:



$$\text{حيث ج} = \text{مجموع} \frac{\text{مربع تكرار كل خلية}}{\text{مجموع التكرارات لعمود تلك الخلية} \times \text{مجموع التكرارات لصف نفس الخلية}}$$

مثال: الجدول الآتي يوضح توزيع عينة من المفحوصين حسب منطقة الإقامة والصحف التي يقرؤونها:

منطقة الإقامة	الصحف المقروءة			المجموع
	الصحف القومية	الصحف الحزبية	الصحف المستقلة	
شعبية	٤	٦	٣	١٣
متوسطة	٣	٢	٤	٩
راقية	٥	٧	٦	١٨
المجموع	١٢	١٥	١٣	٤٠

للحصول على معامل التوافق لا بد من الحصول على القيمة (جـ) وذلك لكل خلية من خلايا الجدول ، ثم جمع قيم (ج) لكل هذه الخلايا، حتى يمكن تطبيق المعادلة السابقة

ففيما يخص الخلية الأولى نجد أنها تتضمن ٤ تكرارات، ومجموع تكرارات العمود هو ١٢ كما أن مجموع تكرارات الصف هو ١٣ وبالتالي فإن (ج) لهذه الخلية يكون

$$\frac{٤ \times ٤}{١٢ \times ١٣} = ٠,١٠٣$$

فيما يخص الخلية الثانية نجد أنها تتضمن ٦ تكرارات، ومجموع تكرارات العمود هو ١٥ كما أن مجموع تكرارات الصف هو ١٣ وبالتالي فإن (ج) لهذه الخلية يكون

$$\frac{٦\times٦}{١٥\times١٣} = ٠,١٨٥$$

وهكذا فيما يخص كل خلية من الخلايا التي تتضمن التكـــرارات موزعـــة علـــى المتغيرين محل الدراسة، فالخلية التاسعة مثلاً تتضمن ٦ تكرارات، ومجموع تكـــرارات العمود هو ١٣ كما أن مجموع تكرارات الصف هو ١٨ وبالتالي فإن (ج) لهذه الخلية يكون

$$\frac{٦\times٦}{١٥\times١٨} = ٠,١٥٤$$

بعد ذلك نقوم بجمع القيم التي حصلنا عليها، لنفرض أن هذه القيم :

٠,١٠٣ + ٠,١٨٥ + ٠,٠٥٣ + ٠,٠٨٣ + ٠,٠٣ + ٠,٠٥١ + ٠,١٢ + ٠,١٨١ + ٠,١٥٤ = ٠,٩٦

أي أن القيمة (ج) تساوي ٠,٩٦ فنحصل بذلك على معامل التوافق كالآتي:

$$ق = \sqrt{١ - \frac{١}{٠,٩٦}} = ٠,٢٠٤$$

ويتم التعرف على الدلالة الإحصائية لمعامل التوافق من خلال حساب قيمة كـــا٢ ، وذلك بموجب المعادلة:

$$\text{كا}^{٢} = \frac{\text{ن} \times \text{ق}^{٢}}{١ - \text{ق}^{٢}}$$

علماً بأن الرمز (ن) هو عدد أفراد العينة، أما الرمز (ق^{٢}) فهو مربع معامل التوافق المحسوب. وفي مثالنا السابق فإن عدد مفردات العينة هو ٤٠ أما معامل التوافق فهو ٠,٢٠٤ وعلى ذلك تكون قيمة كا^{٢} تساوي:

$$١,٧٥٣ = \frac{٤٠ \times ٠,٠٤٢}{١ - ٠,٠٤٢}$$

أي أن قيمة كا^{٢} تساوي ١,٧٥٣ وعلينا أن نقارن هذه القيمة المحسوبة بقيمة كا٢ الجدولية عند درجة حرية تساوي (عدد الأعمدة — ١) × (عدد الصفوف — ١)، ومستوى ثقة ٠,٩٥ أو ٠,٩٩ فإذا كانت قيمة كا^{٢} المحسوبة أقل من قيمة كا٢ الجدولية، فإن ذلك يعني أنه لا توجد علاقة بين الصحف المقروءة ومنطقة الإقامة، أما إذا كانت قيمة كا^{٢} المحسوبة أكبر من قيمة كا^{٢} الجدولية، فإن ذلك يعني أنه توجد علاقة بين الصحف المقروءة ومنطقة الإقامة.

مرة أخرى، فإن البرامج الإحصائية الجاهزة تمكننا من الحصول على هذه المعطيات الإحصائية بسهولة، لكن المعرفة بها أمر أساسي للبحث العلمي. أخيراً، فإنه عند تفسير معامل التوافق، فإنه يجب تذكر أن قيمته دائماً موجبة، كما أن القيمة العظمى لمعامل التوافق لا تصل إلى الواحد الصحيح، لأن هذه القيمة تعتمد على العدد الأصغر للأعمدة أو الصفوف.

٤– معامل كرامير

يعتبر معامل كرامير Cramer coefficient أحد أهم الأساليب الإحصائية التي يمكن استخدامها للكشف عن العلاقة بين المتغيرات الاسمية Nominal، فهذا المعامل يقيس التوافق بين متغيرين اسميين (نوعيين أو كيفيين) وذلك بموجب القانون:

$$\text{ق} = \frac{\text{ج} - ١}{\text{ع} - ١}$$

$$\text{حيث ج} = \text{مجموع} \frac{\text{مربع تكرار كل خلية}}{\text{مجموع التكرارات} \times \text{مجموع التكرارات لعمود تلك الخلية}}$$

ويتم تطبيق هذه المعادلة للحصول على القيمة (جـ) وذلك لكل خلية من خلايا الجدول ، ثم جمع قيم (ج) لكل هذه الخلايا، وذلك على النحو السابق توضيحه في معامل التوافق

ع : العدد الأقل لخلايا الصفوف أو خلايا الأعمدة

مثال توضيحي: لنفرض أننا حصلنا على القيمة (ج) وتبين أنها ١,٢٨ وأن عدد خلايا الأعمدة هو ٤ بينما عدد خلايا الصفوف ٣ في هذه الحالة فإن معامل كرامير يساوي

$$\sqrt{\frac{\text{١,٢٨} - \text{١}}{\text{٣} - \text{١}}} = \text{٠,٣٧٤}$$

غير أنه إذا اشتمل الجدول التكراري المزدوج للبيانات على صفين فقط وعمودين فقط فإن معامل كرامير يتم الحصول عليه من خلال المعادلة:

$$\text{ق} = \sqrt{\text{ج} - \text{١}}$$

فإذا كانت قيمة (ج) تساوي ١,٢٨ فإن معامل كرامير يكون:

$$\sqrt{\text{١,٢٨} - \text{١}} = \text{٠,٥٣}$$

أي أن قيمة معامل كرامير تساوي ٠,٥٣ لكن التعرف على الدلالة الإحصائية لمعامل كرامير يقتضي حساب كا٢ ومقارنة قيمة كا٢ المحسوبة بقيمة كا٢ الجدولية وذلك على النحو السابق توضيحه عند شرح الدلالة الإحصائية لمعامل التوافق

٥- معامل تشيبرو :

سبقت الإشارة إلى أن معامل فاي (ϕ) يتعامل مع متغيرين ينقسم كل منهما انقساماً ثنائياً (نعم – لا)، إلا أنه في بعض الحالات نجد أن أحد المتغيرين ذو انقسام ثلاثي مثل (نعم – متردد – لا) أو أن للبند ثلاث أو أربع استجابات مثل (موافق بشدة – موافق – أرفض – أرفض بشدة)، في مثل هذه الحالات لا يمكن استخدام معامل فاي الذي يعتمد على التقسيم الثنائي للمتغير، وإنما من الممكن استخدام معامل تشيبرو Tachuprou Coefficient والمعروف بمعامل الارتباط الثلاثي ، وقانونه هو:

$$ر_{ش} = \sqrt{\frac{ق^2}{\sqrt{(ل-1)(م-1)}\ (1 - ق^2)}}$$

حيث:

ر ش : معامل تشيبرو

ق : معامل التوافق الذي يمكن حسابه بأي من الطرق السابقة (وبالتالي فإن $ق^2$ تعني مربع معامل التوافق، فإذا كان معامل التوافق الذي حصلنا عليه هو ٠,٤٢ فإن مربع هذه القيمة يساوي ٠,١٧٦٤)

ل: عدد البدائل أو الانقسامات للمتغير الأول (فإذا افترضنا أن المتغير الأول هو قراءة الصحف (يقرأ- لا يقرأ)، فهذا يعني أن عدد بدائل هذا المتغير اثنان (بمعنى استجابتين)

م: عدد البدائل أو الانقسامات للمتغير الثاني ((فإذا افترضنا أن المتغير الثاني هو محل الإقامة (منطقة حضرية- منطقة ريفية- منطقة بدوية)، فهذا يعني أن عدد بدائل هذا المتغير ثلاثة (بمعنى ثلاثة استجابات)

وبهذا تكون كافة عناصر المعادلة معروفة، فإذا افترضنا أننا حصلنا على معامل توافق يساوي ٠,٤٢ بين متغير قراءة الصحف (يقرأ - لا يقرأ) وبين متغير آخر هو محل

الإقامة (منطقة حضرية- منطقة ريفية- منطقة بدوية)، فإن عدد بدائل للمـــتغير الأول يكون ٢ بينما عدد البدائل للمتغير الثاني هو ٣ وبالتالي فإن معامل تشيبرو يكون:

$$\text{ر ش} = \sqrt{\frac{(\text{٠,٤٢})}{\sqrt{(\text{٢}-\text{١})\,(\text{٣}-\text{١})}\;\;(\text{١}-(\text{٠,٤٢}))}}$$

$$\text{ر ش} = \sqrt{\frac{\text{٠,١٧٦٤}}{\sqrt{(\text{١})\,(\text{٢})}\;\;(\text{١}-\text{٠,١٧٦٤})}}$$

$$\text{ر ش} = \sqrt{\frac{\text{٠,١٧٦٤}}{\sqrt{\text{٢}}\times\text{٠,٨٢٣٦}}}$$

$$\text{ر ش} = \sqrt{\frac{\text{٠,١٧٦٤}}{\text{١,١٦١}}}$$

وبما أن ٠,١٧٦٤ ÷ ١,١٦١ = ٠,١٥٢

إذن: $\sqrt{\text{٠,١٥٢}} = \text{٠,٣٩}$

أي أن معامل تشيبرو يساوي ٠,٣٩ وبالتالي فإن قيمة معامل الارتباط بين قـــراءة الصحف ومحل الإقامة يساوي ٠,٣٩ وهذه القيمة هي القيمة المـــصححة لمعامـــل

التوافق الذي سبق أن حصلنا عليه بين هذين المتغيرين والذي كان يبلغ ٠,٤٢ كمـــا سبقت الإشارة

٦- اختبار كا^{٢} للاستقلالية

من المعروف أن كا^{٢} له استخدامات متعددة، من بينها أنه يستخدم كاختبار لمـــدى استقلالية متغيرين Chi Square Test of Independence . إن مدى استقلالية متغيرين يقصد به اختبار ما إذا كانت الظاهرتان مستقلتين – أي غير مرتبطتين أم لا ، ففي مجال الإعلام قد يريد الباحث- على سبيل المثال- معرفة ما إذا كانت آراء المفحوصين (عينة البحث) بشأن الموافقة على فاعلية التليفزيون في المعرفة بتنظيم الأسرة (موافق أو غـــير موافق) تختلف باختلاف خصائص هؤلاء المفحوصين من حيث مكان الإقامة (منطقـــة ريفية، منطقة حضرية، منطقة صحراوية). في هذه الحالة، يتم تطبيق قانون كا^{٢} وهو:

$$\text{كا}^{٢} = \text{مج} \frac{(\text{هـ} - \text{ق})^{٢}}{\text{ق}}$$

حيث هـــ المشاهد (التكرارات الموجودة في كل خلية من خلايا الجدول)

ق المتوقع (التكرارات التي نتوقع أن توجد في كل خلية من خلايا الجدول)، ويـــتم حساب (ق) أي التكرارات المتوقعة بموجب القانون:

$$\text{ق} = \frac{\text{مجموع تكرارات الصف} \times \text{مجموع تكرارات العمود}}{\text{عدد العينة}}$$

وذلك بالنسبة لجميع الخلايا الناتجة عـــن الرصـــد المـــزدوج Cross – Tabulation للبيانات (التكرارات)، مع ملاحظة أن الرمز (ن) هو عدد أفراد العينة الـــتي تم إجـــراء البحث عليها وبدرجات حرية = (عدد الأعمدة – ١) × (عـــدد الـــصفوف – ١). إن الفرض المذكور قد يكون: "تختلف نسب الموافقين على فاعلية التليفزيون في نشر المعرفة

بتنظيم الأسرة باختلاف منطقة الإقامة". ومن خلال تطبيق القانون السابق يتم التعــرف على عدد التكرارات المتوقعة بمعلومية عدد التكرارات المشاهدة بكل خلية مـــن خلايـــا الجدول المزدوج الذي يتضمن توزيع عينة البحث حسب المتغيرين المذكورين، المتغير الأول هو الموافقة على فاعلية التليفزيون في نشر المعرفة بتنظيم الأسرة (أوافق & لا أوافق) أما المتغير الثاني فهو منطقة الإقامة (منطقة ريفية، منطقة حضرية، منطقة صحراوية

مثال: الجدول الآتي يوضح توزيع عينة من المبحوثين (ن= ٢٠٠) حسب محل الإقامة والرأي في فاعلية التليفزيون بشأن تنظيم الأسرة:

المنطقة	الموافقة		المجموع
	موافق	غير موافق	
ريفية	٤٢	١٨	٦٠
حضرية	٣٣	٢٧	٦٠
صحراوية	٣٨	٤٢	٨٠
المجموع	١١٣	٨٧	٢٠٠

في هذا الجدول هناك ٦ خلايا، الخلية الأولى عدد تكراراها ٤٢ أما الخلية الثانية فعـــدد تكراراها ١٨ وفي الخلية الثالثة ٣٣ تكراراً، بينما في الخلية الرابعة يوجد ٢٧ تكراراً، وفي الخلية الخامسة هناك ٣٨ تكراراً أخيراً فإن الخلية السادسة يوجد بها ٤٢ تكراراً. وبموجب القانون السابق ذكره، يتم حساب التكرار المتوقع لكل خلية، على سبيل المثـــال فـــإن التكرار المتوقع للخلية الأولى (التي بها ٤٢ تكرارا) يكون كالآتي:

(ق) الخلية الأولى $\frac{١١٣ \times ٦٠}{٢٠٠} = ٣٣,٩$

(ق) الخلية الثالثة

$$\frac{113 \times 80}{200} = 45.2$$

بهذه الطريقة يتم حساب التكرار المتوقع لكل خلية من خلايا الجدول، فيكون لدينا ست تكرارات متوقعة (بواقع تكرار لكل خلية) ، بعد ذلك نستخدم كل تكرار متوقع لكل خلية حتى نحصل على كا٢ لهذه الخلية. كمثال على ذلك فإن التكرارات الموجودة (المشاهدة) في الخلية الأولى هي ٤٢ تكراراً أما التكرارات المتوقعة والتي تم حسابها لهذه الخلية فتبلغ ٣٣,٩ ويتم حساب قيمة $كا^2$ للخلية الأولى كالآتي:

$$\frac{(42 - 33.9)^2}{33.9} = 0.239$$

أما قيمة كا٢ للخلية الثالثة فيساوي:

$$\frac{(38 - 45.2)^2}{45.2} = 1.15$$

بهذه الطريقة نحصل على قيمة $كا^2$ لكل خلية من خلايا الجدول، فيكون لدينا ست قيم. إن مجمل هذه القيم الست هو $كا^2$ للجدول ككل. وعندما تحسب قيمة $كا^2$ فإن معنوية هذه القيمة تحسب عند درجة حرية معينة، وهذه الدرجة يتم التعرف عليها بموجب القانون:

(عدد الأعمدة – ١) (عدد الصفوف – ١)

وفي مثالنا المذكور فإن عدد الأعمدة هو ٢ بينما عدد الصفوف هو ٣

أي أن درجة الحرية = (٢-١) (٣-١) = ٢

أي أن درجة الحرية التي سنحتكم إليها في معرفة معنوية قيمة $كا^{٢}$ إنما هي تساوي ٢ وفي مثالنا المذكور نفرض أن قيمة $كا^{٢}$ التي حصلنا عليها كانت ٣,٦ فهل هذه القيمة جوهرية أم أنها غير جوهرية وترجع إلى الصدفة؟ وإذا كانت جوهرية (حقيقية) فهل هي جوهرية عند مستوى الثقة ٠,٩٥ أم أنها جوهرية عند مستوى الثقة ٠,٩٩ ؟

لا شك في أن البرامج الإحصائية الجاهزة تتيح لنا هذه المعطيات، غير أن الباحث — في حالة عدم استخدام هذه البرامج - يمكنه استخدام الجداول الإحصائية ويقارن قيمة $كا^{٢}$ التي حصل عليها بقيمة $كا^{٢}$ الموجودة في هذه الجداول عند مستوى الثقة ودرجة الحرية. فإذا كانت قيمة $كا^{٢}$المحسوبة (التي حصل عليها الباحث) أكبر من قيمة $كا^{٢}$ الجدولية، فإن هذا يعني أن المتغيرين محل الدراسة غير مستقلين عن بعضهما البعض، أي أنهما مرتبطان. أما إذا كانت قيمة $كا^{٢}$المحسوبة (التي حصل عليها الباحث) أصغر من قيمة كا٢ الجدولية، فإن هذا يعني أن المتغيرين محل الدراسة مستقلين عن بعضهما البعض، أي أنهما غير مرتبطين.

مرة أخرى فإن $كا^{٢}$ودلالته ، وكذلك كافة المعاملات الإحصائية الأخرى يمكن الحصول عليها باستخدام البرامج الإحصائية الجاهزة ، لكن على الباحث أن يكون لديه المعرفة الكافية بهذه المعاملات وشروط استخدامها وكيفية استخدام المعطيات الكمية لهذه البرامج

٧- معامل لامدا

يستخدم معامل لامدا Lambda Coefficient لقياس الارتباط بين المتغيرات الاسمية المنقسمة، وهو يقيس ما يقيسه $كا^{٢}$، لكن من الأفضل استخدام معامل لامدا عندما تكون بعض تكرارات خلايا الجدول المزدوج بها أصفار. ومن خلال معامل لامدا يمكن معرفة العلاقة بين متغير معين (س) ومتغير آخر (ص) وذلك من خلال القانون:

$$\lambda_{س ص} = \frac{مج\ ك^{١}_{س} - ك^{١}_{ص}}{ن - ك^{١}_{ص}}$$

حيث λس ص = معامل لامدا الذي يوضح الدرجة التي يمكن بها تقدير المـــتغير التابع (ص) من المتغير المستقل (س)

مج ك[1] س : مجموع تكرارات الفئات المنوالية (الأكثر تكراراً) لكل فئة مـــن فئـــات المتغير المستقل س

ك[1]ص : تكرار الفئة المنوالية (الأكثر تكراراً) المجملة للمتغير التابع ص

مثال:

في دراسة موسعة على الشباب الذين يشاهدون نشرات الأخبار كانت العينـــة ٢١٣ مفردة ، و الجدول الآتي يوضح توزيع العينة حسب السن والقضايا محل اهتمامهم، فهـــل هناك ارتباط بين هذين المتغيرين (السن ولقضايا محل الاهتمام)؟

القضايا	السن			المجموع
	٢٠ –٢٥	**٢٦–٣٠**	**٣١– ٣٥**	
سياسية	٣٢	١٣	١٢	٥٧
اقتصادية	٢٤	٣٥	٢٧	٨٦
اجتماعية	١٨	١٠	٤٢	٧٠
المجموع	٧٤	٥٨	٨١	٢١٣

فالمتغير المستقل هو السن (س) ، أما المتغير التابع فهو القضايا محل الاهتمام (ص)

يتم حساب معامل لامدا بالاعتماد على تكرارات الفئات المنوالية للمـــتغير المـــستقل، ومن الجدول نتبين هذه التكرارات: ٣٢، ٣٥، ٤٢ وكذلك على تكرار الفئة المنواليـــة

المجملة للمتغير التابع ، ومن الجدول نتبين أن هذه التكرارات ٧٠ تكراراً. أي أن مـــج ك١ س = ٣٢+ ٣٥+ ٤٢ وبتطبيق المعادلة المذكورة فإن:

$$\lambda_{\text{س ص}} = \frac{(٣٢+٣٥+٤٢) - ٧٠}{٢١٣ - ٧٠} = ٠,٢٧$$

أي أن معامل لامدا يساوي ٠,٢٧ ولمعرفة الدلالة الإحصائية لمعامل لامدا يجب حساب قيمة $كا^٢$ من تكرارات الجدول ثم الكشف عن دلالتها الإحصائية عند درجة حريـــة = (عدد الأعمدة — ١) × (عدد الصفوف — ١)، فإن اتضح أن $كا^٢$ دالة إحصائيا كــــان ذلك معناه أن معامل لامدا له دلالة إحصائية والعكس صحيح. ومن المعروف أن توزيع λ توزيع معقد ومع ذلك فإنه عندما يساوي (الصفر) في المجتمع فإن قيمة $كا^٢$ سوف تكون كذلك مساوية للصفر.

ثانياً: الارتباط بين متغيرين رتبيين

يقصد بالمتغير الرتبي (Ordinal Variable) الخاصية التي يظهر فيهـــا المفحوصـــون في مجموعات متمايزة وفق ترتيب تصاعدي أو تنازلي، لنفرض أن لدينا عينة مـــن ٥٠٠ شخص، وكانت مشاهدتهم اليومية لبرامج التليفزيون كالآتي:

— أقل من ساعة ٢٤٥

— من ساعة إلى أربع ساعات ١٥٦

— أكثر من أربع ساعات ٩٩

فإذا أعطينا تلك المجموعات الثلاث — حسب كثافة مشاهدة التليفزيون – رتباً هي : ١، ٢، ٣ على التوالي، فإن هذا يعني أن المجموعة التي تشاهد التليفزيون أقل من ساعة تأخذ الرتبة (١)، أما المجموعة التي تشاهد التليفزيون من ساعة لأقل من أربع ساعات فتأخذ الرتبة (٢)، في حين أن المجموعة التي تشاهد التليفزيون أكثر من أربع ساعات

تأخذ الرتبة (٣) ومن الواضح أن الرتبة تعكس كثافة المشاهدة، بمعنى أن المجموعة التي تشاهد أكثر أخذت رتبة أعلى، وعلى الرغم من أن أرقــام الرتــب متتاليــة، إلا أن ساعات المشاهدة ليست متساوية البعد عن بعضها البعض، فإذا كان بعض المفحوصين قد تم ترتيبهم في الرتبة (٢) فإن ذلك لا يعني أهم يشاهدون التليفزيون ساعات تعــادل ضعف عدد ساعات المشاهدة الخاصة بالمفحوصين الذين يقعون ضمن الرتبة (١)، كما أن المفحوصين الذين تم ترتيبهم في الرتبة (٣)، لا يعني أهم يشاهدون التليفزيون ثلاثــة أمثال المشاهدة الخاصة بالمفحوصين الذين تم ترتيبهم في الرتبة (١)، بمعنى أكثر تحديداً، فإن المتغير الرتبي يعني مجرد ترتيب المفردات في خاصية معينة ولا يعطي صورة واضحة عن الفروق بين الأفراد في الرتب المختلفة من حيث الخاصية أو الصفة محل القياس. فإذا كان لدينا متغيران رتبيان ونريد معرفة الارتباط بين هذين المتغيرين، فإنه يمكننا استخدام أي من : معامل ارتباط سبيرمان للرتب، معامل جاما، معامل ارتباط كندال، معامــل اتفاق كندال، معامل اتساق كندال، وفيما يلي توضيح لتلك الأساليب

١- معامل ارتباط سبيرمان للرتب:

كثيراً ما يهدف البحث إلى قياس التغير الاقتراني القائم بين ترتيب الأفراد أو الأشياء في صفة وترتيبهم في صفة أخرى ، وهنا يستخدم معامل ارتباط سبيرمان للرتــب Spearman's Coefficient of rank correlation ، كما يمكن اســتخدام هــذا المعامل إذا كان أحد المتغيرين أو كلاهما من النوع الفئوي أو النسبي وذلك بعــد تحويل البيانات إلى رتب.

وفي مجال الإعلام قد يريد الباحث معرفة الارتباط بين ترتيب الأفراد مــن حيــث مشاهدة التليفزيون وقراءة الصحف، وفي كثير من الأحيان يصعب قياس مــتغير مــا قياساً كمياً،ولكن يسهل تعيين رتب للصفة أو السلوك المراد دراسته عن هذا المتغير، فمثلاً إذا كان لدينا خمسة مقالات صحفية وأردنا التمييز بين هذه المقــالات مــن حيث إعجاب القراء بها ، نجد أنه يسهل على الأفراد ترتيب تلك المقالات حسب درجة إعجابهم بها ، كأن نطلب من المبحوثين إعطاء تقدير لكل مقال (ضعيف جداً، ضعيف، متوسط، أعلى من المتوسط، ممتاز) وتكون الرتب لهذه التقــديرات علــى

التوالي هي (١، ٢، ٣، ٤، ٥) ، وقد يريد الباحث معرفة ارتباط ذلك بمستوى التعليم (يقرأ ويكتب، الابتدائية، الإعدادية، الثانوية، جامعي فأعلى)، وتكون الرتب لهذه المستويات على التوالي هي (١، ٢، ٣، ٤، ٥) ، أي أن كلا المتغيرين يأخذ رتباً معينة، وهنا يمكن استخدام معامل ارتباط سبيرمان للرتب وفق القانون:

$$\text{مـ} = ١ - \frac{٦ \text{ مج ف}^{٢}}{\text{ن (ن}^{٢} - ١)}$$

حيث مـ : معامل الارتباط

ن : عدد أزواج البيانات (للمتغيرين) أو عدد أفراد العينة.

ف : الفرق بين رتب المتغيرين.

فإذا كان لدينا عينة قوامها ٣٠ من الشباب، وقد حصل كل منهم على درجة معينة على مقياس الاستخدامات والإشباعات، ودرجة أخرى على مقياس المعرفة، فإننا نقوم بالآتي:

- ترتيب درجات هؤلاء المفحوصين على المقياس الأول، وكذلك ترتيب درجاتهم على المقياس الثاني، فيكون لكل مفحوص ترتيبان (الترتيب على المقياس الأول والترتيب على المقياس الثاني)
- حساب الفروق بين ترتيب كل مفحوص في الدرجة على المقياس الأول وترتيب المفحوص نفسه في الدرجة على المقياس الثاني، فإذا كان أحد المفحوصين قد حصل على الترتيب السابع في الدرجة على المقياس الأول وحصل على الترتيب العاشر في الدرجة على المقياس الثاني، فإن الفرق يكون ٧- ١٠ = -٣
- تربيع الفروق في الترتيب وذلك لكل مبحوث في العينة،،، فإذا كان الفرق بين الترتيبين للمبحوث الأول هو -٣ فإن تربيع هذا الفرق يكون : -٣ × -٣ = ٩ وهكذا لكل فرد من أفراد العينة
- جمع مربعات الفروق، فنحصل بذلك على مج ف٢

وباستخدام المعادلة المذكورة يمكن حساب معامل ارتباط الرتب، فإذا كانت العينة

تتكون من ٣٠ شخص وكان مجموع مربعات الفروق هو ٤٢٨ مثلاً فإن معامل ارتباط الرتب:

$$\text{م} = ١ - \frac{٦ \ (٤٢٨)}{٣٠ \ (٩٠٠ - ١)} = ٠,٠٩٥$$

ويفضل استخدام معامل ارتباط الرتب لسبيرمان في حالة العينات التي يكون حجمها ١٠ فأقل ومن الممكن استخدامه بوجه خاص حينما لا يتجاوز حجم العينة ٣٠ فرداً. وعند ترتيب المتغيرين المراد تعيين معامل الارتباط بينهما يجب أن يتم الترتيب من الأكبر إلى الأصغر بنفس الطريقة للمتغيرين معاً، أو من الأصغر إلى الأكبر بنفس الطريقة للمتغيرين معاً ولا يصح ترتيب أحد المتغيرين من الأكبر إلى الأصغر وترتيب المتغير الآخر من الأصغر إلى الأكبر.

٢– معامل جاما :

وهو يستخدم في حالة المتغيرات الرتبية، ويشيع استخدامه بكثرة إذا كان لدينا رتبتان فقط لكل متغير، خاصة عندما تتكرر بعض القيم في المتغير الواحد، ويطلق على البيانات من هذا النوع البيانات المتساوية (ذات الصلة) Tied data وحينما يصاحب ذلك عدد كبير لأزواج قيم المتغيرين. فقد يجد الباحث مثلاً أن تصنيف المفحوصين حسب مشاهدة التليفزيون والدخل يتخذ رتبتين فقط لكل متغير، فحسب مشاهدة التليفزيون يصنف المبحوثون إلى (متوسط، مرتفع) ، وذلك بقيم ترتيبية (١ ،٢) على التوالي، أما حسب الدخل فإن تصنيف المفحوصين يتخذ هو الآخر رتبتين (منخفض، متوسط) وذلك بقيم ترتيبية (١ ،٢) على التوالي، في هذه الحالة يكون من المناسب استخدام معامل ارتباط جاما (Gamma)، لأن الترتيب لكل متغير له نطاق ضيق مما يؤدي إلى زيادة في التكرارات لنفس الرتبة ولنفس الأشخاص، ويعتمد قانون معامل جاما على جدول التكرار المزدوج وعلى حالات الاتفاق و الاختلاف بين أزواج القيم، ويتمثل هذا القانون في:

$$ج = \frac{ن_{ق} - ن_{خ}}{ن_{ق} + ن_{خ}}$$

حيث ج: معامل جاما

$ن_{ق}$: حاصل ضرب عدد حالات الاتفاق

$ن_{خ}$:حاصل ضرب عدد حالات الاختلاف

مثال: الجدول الآتي يوضح توزيع عينة من المبحوثين حسب معدل قراءة الصحف ومعدل مشاهدة التليفزيون

المشاهدة	معدل قراءة الصحف	
	منخفض	متوسط
منخفض	٢٨ (أ)	٤١ (ب)
متوسط	١٢ (ج)	٢٣ (د)

إن معامل جاما يعتمد على تكرارات الاتفاق وتكرارات الاختلاف بين المتغيرين في الجدول المزدوج، ولتوضيح ذلك، فإن الخلية الأولى (ولنرمز لها بالخلية أ) تتضمن ٢٨ تكراراً ، أما الخلية الثانية(ولنرمز لها بالخلية ب فإنها تتضمن ٤١ تكراراً في حين تتضمن الخلية الثالثة (الخلية ج) ١٢ تكراراً، وأخيراً فإن الخلية الرابعة (الخلية د) تتضمن ٢٣ تكراراً ، ونتبين من الجدول أن هناك ٢٨ مبحوثاً من ذوي المستوى المنخفض من حيث مشاهدة التليفزيون ومعدل قراءة الصحف (منخفض & منخفض) ، كما أن هناك ٤١ مبحوثاً من ذوي المستوى المنخفض من حيث مشاهدة التليفزيون وفي الوقت نفسه لديهم مستوى متوسط من حيث معدل قراءة الصحف (منخفض & متوسط) ،،،

وهكذا.

فإذا رجعنا إلى معادلة معامل جاما، فإن الرمز $ن_{ق}$ هو حاصل ضرب عدد حالات الاتفاق، ونحصل عليه بضرب تكرارات الخلية (ا) في تكرارات الخلية (د)، أي ٢٨× ٢٣= ٦٤٤ أما الرمز $ن_{خ}$ فهو حاصل ضرب عدد حالات الاختلاف، ونحصل عليه بضرب تكرارات الخلية (ب) في تكرارات الخلية (ج)، أي ٤١× ١٢= ٤٩٢ فإن قيمة جاما تساوي

$$ج = \frac{٦٤٤ - ٤٩٢}{٦٤٤ + ٤٩٢} = ٠,١٣٤$$

غير أنه من الجدير بالذكر أن معامل (جاما) يستخدم أيضاً في التصنيفات الأكثر من ثنائية للمتغيرات، كأن نريد إيجاد العلاقة بين معدل مشاهدة البرامج السياسية في التليفزيون (منخفض، متوسط، مرتفع) ومستوى المعرفة بالقضايا الدولية (منخفض، متوسط، مرتفع)، وهنا يكون الترتيب لكل متغير له نطاق ضيق مما يؤدي إلى زيادة في التكرارات لنفس الرتبة ولنفس الأشخاص

٣- معامل ارتباط كندال:

يهدف معامل ارتباط كندال Kendal correlation coefficient إلى قياس العلاقة بين متغيرين كلاهما من النوع الرتبي ويعتمد على نفس فكرة معامل جاما، ويرمز لمعامل ارتباط كندال بالرمز Ta (تو أ) وتقرأ (تو ألف) والقانون المستخدم لذلك على الصورة:

$$تو\ أ = \frac{ن_{ق} - ن_{خ}}{٠,٥\ ن\ (ن-١)}$$

حيث: تو أ معامل ارتباط كندال من النوع أ

كذلك $ن_{ق}$ ، $ن_{خ}$ كما هو في حالة معامل جاما (بمعنى أن: ($ن_{ق}$) حاصل ضرب عدد

حالات الاتفاق ، بينما ($ن_{خ}$) هي حاصل ضرب عدد حالات الاختلاف

أما الحرف (ن) فهو عدد أفراد عينة الدراسة

فإذا كانت $ن_{ق}$ = ٢٥ بينما $ن_{خ}$= ٢٣ وكان عدد مفردات العينة ١٠ فإن معامل ارتباط كندال يساوي:

$$تو\ أ = \frac{٢٥ - ٢٣}{٠,٥ \times ١٠ \times ٩} = \frac{٢}{٤٥} = ٠,٠٤$$

علماً بأنه في حالة وجود قيم تتساوي لها الرتبة أو تتكرر فإن قيمة المعامل لا تصل إلى الحد الأقصى أو ما نسميه الارتباط التام ±١ ويعتبر هذا من المآخذ على معامل ارتباط كندال (من النوع الأول)، لأن هناك أكثر من قيمة يكون لها نفس الرتبة،ومن ثم يتم إجراء تصحيح لتلك القيمة باستخدام معامل ارتباط كندال من النوع الثاني، والذي تتضمنه البرامج الإحصائية تحت مسمى Kendall's tau-b لكن الأهم من ذلك هو أنه عندما يكون عدد الأعمدة لا يساوي عدد الصفوف فإنه في هذه الحالة يجب استخدام معامل ارتباط كندال من النوع الثالث Kendall's tau-c من أمثلة ذلك عندما نريد إيجاد الارتباط بين متغيرين رتبيين الأول هو كثافة مشاهدة التليفزيون (منخفضة، متوسطة، مرتفعة) وتأخذ الرتب ١، ٢، ٣ على التوالي ، والثاني هو الدخل الشهري (٥٠٠ دولار أمريكي فأقل، أكثر من ٥٠٠ دولار أمريكي) وتأخذ الرتب ١، ٢ على التوالي. هنا نلاحظ أن متغير مشاهدة التليفزيون يضم ثلاث فئات أما متغير الدخل فيضم فئتين فقط، أي أن المتغيرين غير متساويين من حيث عدد الفئات أو المجموعات، وبالتالي لا يكون عدد الأعمدة مساوياً لعدد الصفوف:

مشاهدة التليفزيون	الدخل الشهري بالدولار الأمريكي	
	أقل من ٥٠٠ دولار	٥٠٠ دولار فأكثر
منخفضة		
متوسطة		
مرتفعة		

من الواضح أن لدينا ثلاثة صفوف تعبر عن متغير مشاهدة التليفزيون (منخفضة، متوسطة، مرتفعة) مقابل عمودين فقط بشأن متغير الدخل الشهري (٥٠٠ دولار أمريكي فأقل، أكثر من ٥٠٠ دولار أمريكي) في هذه الحالة يجب استخدام معامل ارتباط كندال من النوع الثالث Kendall's tau-c

٤- معامل اتفاق كندال:

يستخدم معامل اتفاق كندال Kendall coefficient of concordance (W) لحساب معامل الاتفاق بين الرتب ، فقد يقتضي البحث حساب الارتباط بين أكثر من ترتيبين كما سبق في حالة معامل ارتباط سبيرمان للرتب، ويتطلب ذلك حساب ارتباط رتب المتغير الأول برتب المتغير الثاني ثم ارتباط رتب المتغير الأول برتب المتغير الثالث ثم ارتباط رتب المتغير الأول برتب المتغير الرابع......ثم ارتباط رتب المتغير الثاني برتب المتغير الثالث، ثم ارتباط رتب المتغير الثاني برتب المتغير الرابع.... وهكذا، على أن يتم تعيين متوسط معاملات الارتباط الناتجة عن ذلك. هذا الأمر بالطبع يحتاج لمزيد من الجهد في المعالجات الحسابية فضلا عن الوقت.

لنفرض أننا عرضنا استبانة تتضمن عدداً من البنود على مجموعة من المحكمين بهدف الكشف عن ترتيبهم لهذه البنود من حيث الأهمية والدلالة للبحث ، ومعرفة مدى اتفاق هؤلاء المحكمين... في هذه الحالة يستخدم معامل اتفاق كندال لأنه يسهل الإجراءات، ويناسب التعرف على درجة الاتفاق بين الرتب، ويتمثل قانون اتفاق كندال في الصورة الآتية:

$$ر = \frac{٢}{م^{٢}} \times \frac{٦ \text{ مج } ف^{٢}}{ن(ن^{٢} - ١)}$$

حيث م : عدد المحكمين

ن : حجم العينة (عدد المحكوم عليهم بنودا كانت أم أفراداً......)

$ف^{٢}$: مربعات الفروق بين مجموع رتب كل مفردة وبين المتوسط العام لمجموع الرتب

مثال: في سياق التخطيط السياسي للتعامل مع قضية الملف النووي الإيراني أعدت وكالة المخابرات المركزية الأمريكية CIA استراتيجية دعائية وإعلامية تتضمن ٥ خطط متكاملة لتنفيذها على المستوى المحلي والدولي حيال هذه المسألة، وتم عرض هذه الإستراتيجية على ثلاثة خبراء. ولإقرار هذه الإستراتيجية لا بد أن يكون معامل الاتفاق بين هؤلاء الخبراء لا يقل عن ٩٠% . بافتراض أن مج ف ٢ يساوي ٧٢ فهل يمكن إقرار هذه الإستراتيجية؟

في هذا المثال نلاحظ أن (م) = ٣ بينما قيمة ن = ٥ ، أما مج ف٢ فهو ٧٢ وبالتالي فإن معامل اتفاق كندال يساوي:

$$ر = \frac{٢}{٩} \times \frac{٤٣٢}{٥(٢٤)} = ٠,٧٩٢$$

أي أن نسبة الاتفاق بين الخبراء حول ترتيب أهمية عناصر الإستراتيجية الدعائية الأمريكية بلغت ٧٩% أي أنها تقل عن ٩٠% كنسبة اتفاق مطلوبة بين الخبراء، وبالتالي لن يتم إقرار الإستراتيجية المذكورة.

وكثيراً ما يستخدم معامل اتفاق كندال في بحوث الاتصال ونحن بصدد التعرف على العلاقة بين ترتيب المفحوصين للبرامج والمواد الإعلامية من حيث أهميتها بالنسبة لهم باعتبار هؤلاء المفحوصين من مستخدمي وسائل الإعلام، كما يطبق معامل اتفاق كندال على استجابات القائمين بالاتصال ونحن بصدد رصد وتحليل آرائهم وتقييماتهم بشأن

قضايا معينة تتخذ رتباً من حيث الأهمية، ..... الخ ، هذا بالإضافة إلى استخدام معامـــل اتفاق كندال في التحقق من صدق أدوات جمع البيانات

٥- معامل اتساق كندال

يستخدم معامل اتساق كندال Kendall coefficient of consistence إذا كانـــت البيانات مرتبة في صورة أزواج pairs، كل زوج يتكون من عنصرين مـــرتبين علـــى أفضل وأقل تفضيل تبعا لمحك معين، ويعرف ذلك بالمقارنات الثنائيـــة أو مقارنــــات الأزواج Paired comparisons

فإذا افترضنا أن لدينا ثلاثة برامج أ، ب، ج عرضت على عينة من المفحوصين وكــــان المطلوب أن يرتب المفحوصون هذه البرامج حسب تفضيلهم لها. في هذه الحالة يكون من الطبيعي أن تختلف درجة تفضيل تلك البرامج، فإذا كان بعض المفحوصين أفادوا أنهم يفضلون البرنامج (أ) على البرنامج (ب) وكذلك يفضلون البرنامج (ب) على البرنامج (ج)، فإن استجابتهم لكي تكون متسقة يجب أن يفضلوا البرنامج (أ) على البرنامج (ج)، أما إذا كانت استجابتهم تفيد أن البرنامج (ج) أفضل من البرنامج (أ) فإن تلك الاستجابات تكون غير متسقة ، وتقل الثقة في معنى الرتب التي يقـــدرها المفحوص كلما زاد عدم الاتساق (Inconsistent) . وعندما تتوافر لدينا مقارنــــات ثنائية أو مقارنات زوجية (Paired)، فإنه بالإمكان إيجاد عدد الاســـتجابات غـــير المتسقة والاستفادة منها في الكشف عن معامل الاتساق في هذه الأحكـــام وذلـــك باستخدام معامل اتساق كندال بموجب المعادلة:

$$\text{ك} = ٢ \times \frac{\text{٦مج ف}^{٢}}{\text{ن} (\text{ن}^{٢} - ١)}$$

فإذا افترضنا أن لدينا استبياناً مكوناً من أسئلة متعددة من بينها خمسة أسئلة لقياس اتساق إجابات المفحوصين فإن ذلك يعني أن قيمة (ن) تساوي ٥   وبالتالي يكـــون لدينا ١٠ مقارنات زوجية (وذلك من خلال : ن (ن-١) ÷ ٢ .بمعنى ٥×٤ ÷ ٢= ١٠   فإذا   كنا نريد معرفة نسبة المقارنات الثنائية المتسقة، فإننا نعطـــي الـــرقم (صفر) لحالات عدم الاتساق والرقم (١) لحالات الاتساق، ولتوضيح هذه الفكـــرة

فإن رصد حالات الاتساق يكون في جدول يتخذ الشكل الآتي:

السؤال	الأول	الثاني	الثالث	الرابع	الخامس	**مجموع**	ف	ف٢
الأول	–	١	صفر	١	١	٣	١,٦	٢,٥٦
الثاني		–	١	١	صفر	٢	٠,٦	٠,٣٦
الثالث			–	صفر	١	١	–٠,٤	٠,١٦
الرابع				–	١	١	–٠,٤	٠,١٦
الخامس					–			

نتبين من الجدول أن إجمالي مجموع رتب جميع الأسئلة يساوي ٧ ونظراً لأن لدينا خمسة أسئلة فإن المتوسط العام يكون ٧÷ ٥ = ١,٤

وللحصول على (ف) فإننا نطرح رتب كل سؤال من قيمة المتوسط العام، ففي السؤال الأول مثلاً تكون قيمة ف= ٣–١,٤ = ١,٦ وهكذا ، وبتربيع قيم (ف) نحصل على $ف^{٢}$ حسبما يتضح من العمود الأخير، وبجمع قيم ف٢ لكل الأسئلة نتبين أن مجمل $ف^{٢}$ يساوي ٣,٢٤ وبهذا يكون معامل اتساق كندال بموجب المعادلة:

$$\text{ك} = ٢ \times \frac{\text{٦مج ف٢}}{\text{ن (ن٢ –١)}}$$

أي أن :

$$\text{ك} = ٢ \times \frac{\text{٦×٣,٢٤}}{\text{٥ (٢٥ –١)}} = ٠,٣٢٤$$

معنى ذلك أن هناك اتساقاً بين ثلث المقارنات الثنائية. إن هذا مثال مبسط هدفه مجرد التوضيح، لأن المقارنات الثنائية في الواقع تكون بالمئات أو بالألوف، ويتم

التعامل معها آلياً باستخدام البرامج الإحصائية الجاهزة، ويمكن الاستفادة من معامـــل اتساق كندال في تطبيقات إحصائية متعددة في البحث العلمي بما في ذلك الكشف عن مدى اتساق استجابات المفحوصين، وكذلك اتساق آراء المحكمين وغير ذلك.

ثالثاً: الارتباط بين متغيرين
أحدهما رتبي والآخر اسمي

إذا كان لدينا متغيرين أحدهما من النوع الرتبي Ordinal والآخر من النوع الاسمي Nominal ونريد رصد معامل الارتباط بينهما، فإننا يمكننا معامل الارتباط الثنائي (معامل كوريتون) وكذلك معامل ثيتا (معامل فريمان)، وذلك على النحو الآتي:

٣-معامل الارتباط الثنائي للرتب:

يستخدم معامل الارتباط الثنائي Rank biserial- أو معامل كوريتون (Coreton Coefficient) عندما نكون بصدد بحث العلاقة بين أحدهما رتبي Ordinal والآخر اسمي Nominal ، فقد يكون لدى الباحث بنود تقيس مستوى تفضيل مشاهدة برامج التليفزيون، ويتخذ مستوى التفضيل رتباً هي: ١، ٢، ٣، ٤، ٥ بحيث تدل الرتبة الأعلى على تفضيل أكبر، ويريد الباحث معرفة قيمة الارتباط بين مستوى التفضيل ومتغير اسمي آخر مثل الجنس (ذكور & إناث) . في هذه الحالة يمكن استخدام الارتباط الثنائي للرتب أو معامل كوريتون بين الجنس ومستوى التفضيل . إن القانون الأساسي لهذا المعامل هو:

$$ر = \frac{٢}{ن} \left[\overline{ص_{٢}} - \overline{ص_{1}} \right]$$

حيث $\overline{ص_{١}}$: متوسط رتب استجابات الأفراد الذكور

$\overline{ص_{٢}}$: متوسط رتب استجابات الأفراد الإناث

ن : عدد أفراد العينة

لتطبيق هذا القانون فإننا نرصد أمام كل فرد في العينة رتبته من حيث النوع أو الجنس (بمعنى أن كل فرد من الذكور يأخذ الترتيب ١ بينما يأخذ كل شخص من الإناث الترتيب ٢) كما نرصد أمام كل فرد في العينة رتبته من حيث تفضيل البرامج التليفزيونية (١ أو ٢ أو ٣ أو ٤ أو ٥) وبهذا يكون لكل فرد ترتيب من حيث الجنس، وترتيب من حيث تفضيل برامج التليفزيون.

ثم نقوم بحساب متوسط رتب مجموعة الذكور من حيث تفضيل البرامج(وذلك بجمع هذه رتب هذه المجموعة من حيث تفضيل البرامج وقسمة الناتج على عدد هذه المجموعة)، المنطق نفسه فيما يخص مجموعة الحضر. وبهذا يكون لدينا متوسط رتب مجموعة الذكور، ومتوسط رتب مجموعة الإناث .

فإذا افترضنا أن العينة تتكون من ٣٠ مفردة (١٠ ذكور ، ٢٠ إناث)، وكان متوسط رتب مجموعة الذكور هو= ١,٦ بينما كان متوسط رتب مجموعة الإناث هو ٨,٣ فإن معامل الارتباط الثنائي يكون:

$$ر = \frac{٢}{٣٠} (٨,٣-١,٦)$$

$$= ٠,٠٧ \times ٦,٧ = ٠,٤٧$$

أي أن معامل الارتباط بين النوع ومستوى تفضيل برامج التليفزيون هو ٠,٤٧

وتحسب الدلالة الإحصائية لمعامل الارتباط الثنائي للرتب من خلال تحويل قيمة معامل الارتباط إلى درجة معيارية (Z)، وذلك بموجب المعادلة:

$$= ر \times ط \times ن \times \sqrt{\frac{ن}{ن_{١} ن_{٢}}}$$

علماً بأن ر: معامل الارتباط الثنائي للرتب

ن١ : عدد الأفراد الذكور

ن٢ : عدد الأفراد الإناث

ن : $ن_1 + ن_2$

ط : طول ارتفاع المنحنى الطبيعي عند النقطة التي تفصل بين النسبة $\frac{ن_1}{ن}$ والنسبة $\frac{ن_2}{ن}$ وذلك باستخدام إحدى هاتين النسبتين (مساحة صغرى أو مساحة كبرى) والكشف في جدول ارتفاعات المنحنى الطبيعي. ففي مثالنا المذكور نلاحظ أن عدد الأفراد الذكور هو ١٠ بينما عدد الأفراد الإناث هو ٢٠ وبذلك فإن :

$$\frac{ن١}{ن} = ١٠ \div ٣٠ = ٠,٣٣٣$$

$$\frac{ن٢}{ن} = ٢٠ \div ٣٠ = ٠,٦٦٧$$

أي أن المساحة الصغرى تساوى ٠,٣٣٣ أما المساحة الكبرى فهي ٠,٦٦٧ ويمكننا استخدام أي من المساحتين لرصد ارتفاع المنحنى الطبيعي (أي أننا نستخدم جدول ارتفاعات هذا المنحنى، وهو أحد الجداول الإحصائية الهامة).. فإذا استخدمنا المساحة الصغرى وهي ٠,٣٣٣ فإننا نجد أن ارتفاع المنحنى المقابل لها هو ٠,٢٤٩٢

وتحسب الدلالة الإحصائية لمعامل الارتباط الثنائي للرتب من خلال تحويل قيمة معامل الارتباط إلى درجة معيارية (Z)، وذلك بموجب المعادلة المذكورة، أي أن :

$$Z = ٠,٤٧ \times ٠,٢٤٩٢ \times ٣٠ \sqrt{\frac{٣٠}{٢٠٠}} = ١,٤$$

أي أن قيمة Z تساوي ١,٤ وهذه القيمة اقل من ١,٩٦ وبالتالي لا توجد علاقة بين النوع ومستوى تفضيل برامج التليفزيون، وذلك عند مستوى المعنوية ٠,٠٥

٢- معامل ثيتا (معامل فريمان)

يستخدم معامل ثيتــا Theta coefficient أو معامــل فريمــان (Freeman coefficient إذا كانت البيانات تتعلق بمتغيرين أحدهما اسمي،الجنس مثلاً (ذكور & إناث) أو الجنسية (مصري & غير مصري) والمتغير الآخر في صورة رتبة (أشــاهد التليفزيون يومياً، أشاهد التليفزيون معظم أيام الأسبوع، أشاهد التليفزيون بعــض أيام الأسبوع، أشاهد التليفزيون يوم واحد في الأسبوع، لا أشــاهد التليفزيــون) ولهذه الاستجابات أرقام رتبية (٥، ٤، ٣، ٢، ١) ويجب أن تكون البيانات عند استخدام هذا المعامل موضحة تميز تكرارات المتغير الاسمي في كل مستوى رتبي، بمعنى أننا إذا كنا بصدد بحث العلاقة بين الجنس (ذكور – إناث) ومشاهدة التليفزيون وفق الرتب المذكورة، فإنه يجب تميز الذكور على الإناث أو العكس في كل مستوى رتبي، كأن يتميز الذكور عن الإناث في المستوى الرتبي الأول، وتتميز الإناث عن الــذكور في المستوى الرتبي الرابع ..... وهكذا في بقية المستويات الرتبية، والتميز هنا يقصد به عدد المفردات، فإذا كان عدد الإناث أكبر من عدد الذكور في مستوى رتبي معين، فإن ذلك يعني تميز الإناث على الذكور في هذا المــستوى، ولا يــشترط أن تكــون البيانات الرتبية متسلسلة، ويرمز لمعامل ثيتا بالرمز اللاتيني Ǿ، وكمثال علــى هــذا المعامل تفرض أننا بصدد بحث معامل الارتباط بين النوع (ذكور & إناث) وترتيب الدرجات في مادة الإحصاء (١، ٢، ٣، ٤، ٥)

$$\text{ثيتا} = \frac{\text{ك ذ ث} - \text{ك ث ذ}}{\text{ن ذ} \times \text{ن ث}}$$

حيث إن:

ك ذ ث : عدد التكرارات التي يكون فيها الذكور في رتبة أعلى من الإناث

ك ث ذ : عدد التكرارات التي يكون فيها الإناث في رتبة أعلى من الذكور

ن ذ : عدد مجموعة الذكور

ن ث: عدد مجموعة الإناث

كمثال توضيحي نفرض أن دراسة أجريت على عينة قوامها ٢٠٠ مفردات من الجنسين (١٦٠ ذكور ، ٤٠ إناث)، وعند تنظيم البيانات تبين أن استجابات الإناث عكست أنهن أتين بواقع ٣٠ تكراراً في رتبة أعلى من الذكور، وذلك من حيث معدل مشاهدة برامج الشباب في التليفزيون. في حين جاء الذكور بواقع ١٢٠ تكراراً في مرتبة أعلى من الإناث من حيث المعدل نفسه. في هذه الحالة فإن (ك ث ذ) تساوي ٣٠ أما (ك ذ ث) فتساوي ١٢٠ وبالتالي فإن معامل ثيتا يكون:

$$\text{ثيتا} = \frac{\text{١٢٠} - \text{٣٠}}{\text{١٦٠} \times \text{٤٠}} = \text{٠,٠١٤}$$

أي أن معامل ثيتا يساوي ٠,٠١٤ وهو ارتباط ضعيف كما هو واضح.

رابعاً:الارتباط بين متغيرين من النوع الفتري (معامل ارتباط بيرسون)

لحساب الارتباط بين متغيرين كلاهما من النوع الفتري يستخدم معامل ارتباط بيرسون (مثال ذلك عندما نريد حساب الارتباط بين الوقت المنقضي في قراءة الصحف والوقت المنقضي في مشاهدة التليفزيون) ويعتبر معامل ارتباط بيرسون مقياساً معيارياً للعلاقة، بمعنى أنه يدخل في حسابه المتوسط والانحراف المعياري لكل من مجموعتي الدرجات المراد إيجاد العلاقة بينهما، وهذا يعني أن أي تحويل خطي لإحدى مجموعتي الدرجات لا يؤثر في قيمة معامل ارتباط بيرسون، وبذلك لا يكون لوحدة القياس أهمية تذكر عند إيجاد معامل الارتباط. ويعد معامل ارتباط بيرسون أكثر أنواع معاملات الارتباط استخداماً في البحوث بمجالاتها المختلفة، بل إن الكثير من أنواع معاملات الارتباط والاقتران الأخرى تعتبر حالات خاصة من معامل ارتباط بيرسون. ولكي يتضح معنى ارتباط بيرسون ربما يكون من الأفضل التعبير عن المتغيرات في صورة درجات معيارية حتى يمكن الربط بين معامل الارتباط وغيره من المقاييس الإحصائية المختلفة. فإذا افتراضنا أن (س)، (ص) تمثل أزواجا من الملاحظات انحرافاتها المعيارية، ($ع_{س}$)، ($ع_{ص}$) على الترتيب، فإنه لتحويل الملاحظات (س)، (ص) إلى درجات معيارية تستخدم الصيغتان التاليتان:

$$د_{س} = \frac{س - \overline{س}}{ع_{س}}$$

$$،\ د_{ص} = \frac{ص - \overline{ص}}{ع_{ص}}$$

وهذه الدرجات المعيارية متوسطها صفر، وانحرافها الواحد الصحيح.

ويمكن تعريف معامل ارتباط بيرسون والذي سنرمز له بالرمز (ر) بأنه متوسط مجموع حاصل ضرب الدرجات المعيارية المتقابلة للمتغيرين (س)، (ص) ويمكن التعبير عن هذا رياضياً بالصيغة التالية:

$$ر = \frac{مجـ (د_{س} \times د_{ص})}{ن}$$

ولذلك فإنه يمكن الحصول على معامل ارتباط بيرسون بين متغيرين (س)، (ص) بتحويل كل قيمة من قيم المتغيرين إلى درجات معيارية باستخدام الصيغتين السابقتين وجمع حاصل ضرب الدرجات المعيارية المتقابلة للمتغيرين، وقسمة الناتج على عدد القيم. ولتوضيح معني الصيغة الرياضية المستخدمة في إيجاد معامل ارتباط بيرسون نفترض أن لدينا أزواجاً من الملاحظات محولة إلى درجات معيارية، فمجموع حاصل ضرب الدرجات المعيارية المتقابلة مج ($د_{س} \times د_{ص}$) مقسوماً على عدد المفردات يعد مقياسا لدرجة العلاقة بين المتغيرين. وتصل مج ($د_{س} \times د_{ص}$) إلى قيمتها العظمي:

(١) إذا كانت قيم ($د_{س}$)، ($د_{ص}$) لها نفس الترتيب

(٢) وإذا ساوت كل قيمة من قيم ($د_{س}$) القيمة المناظرة لها ($د_{ص}$)، أي إذا تساوت قيمة مجموعتي الملاحظات

ويستند معامل ارتباط بيرسون إلى افتراضات يجب أن يتحقق منها الباحث في المتغيرات التي يود دراسة العلاقة بينها، وتتلخص تلك الافتراضات في أن معامل ارتباط بيرسون هو

مقياس للعلاقة الخطية أو المستقيمة بين متغيرين، ويمكن للباحث التحقق مبدئياً من استقامة العلاقة برسم الشكل الانتشاري لقيم المتغيرين وتأمل الشكل الناتج، فإذا اتضح للباحث أن توزيع القيم يتخذ شكلاً بيضاوياً دون أي نزعة انحناء فإن هذا يمكن أن يكون دليلاً على استقامة العلاقة، وإن كان ابتعاد العلاقة ابتعاداً طفيفاً عن الاستقامة لا يمنع الباحث من استخدام معامل ارتباط بيرسون كتقريب مبدئي لقيم معاملات الارتباط الأخرى التي يمكن أن يستخدمها في حالة العلاقة غير المستقيمة. أما إذا ابتعد شكل العلاقة عن الاستقامة وأصبح واضحاً للباحث من تأمله للشكل الانتشاري أن العلاقة بين المتغيرين منحنية، فإنه يجب أن يستخدم ما يسمي بنسبة الارتباط Correlation Ratio، أو أي أسلوب إحصائي آخر يتفق وهذه العلاقة المنحنية

ويمكن حساب معامل ارتباط بيرسون للبيانات غير المبوبة (سواء من خلال استخدام الدرجات المعيارية، أو باستخدام الانحرافات عن المتوسط ، أو باستخدام الدرجات الخام مباشرة، أو باستخدام الفروق بين الدرجات الخام) ، كما يمكن حساب معامل ارتباط بيرسون للبيانات المبوبة، وتبدو أهمية ذلك عندما تشتمل البيانات على عدد كبير من أزواج القيم، حيث يمكن تبويب (جدولة) تلك القيم في جدول تكراري مزدوج Two-way frequency table ثم إيجاد معامل ارتباط بيرسون لهذه البيانات المبوبة باستخدام طريقة الترميز Code method . غير أن استخدام تلك الطريقة يتطلب تصحيح معامل الارتباط من الأخطاء الناتجة عن تبويب البيانات، ذلك أن قيمة معامل الارتباط من البيانات المبوبة تكون قيمة تقريبية، والسبب في ذلك يرجع إلى أننا اعتبرنا أن تكرار كل فئة يقع في مركز تلك الفئة وكلما زاد طول الفئة زاد بالطبع الناتج عن هذا التقريب. فإذا أراد الباحث أن يحصل على القيمة المضبوطة لمعامل ارتباط بيرسون فعليه أن يستخدم الدرجات الخام مباشرة بدلاً من استخدام طريقة الترميز

أما إذا استخدم الباحث طريقة الترميز وكان عدد فئات أي من المتغيرين قليلاً فإن تقدير قيمة معامل الارتباط تكون أقل مما لو استخدم طريقة الدرجات الخام. وفي الحالات المتطرفة التي يكون فيها عدد فئات أي من المتغيرين فئتين فقط تقل قيمة معامل الارتباط الناتجة عن استخدام طريقة الترميز بقدر ثلثي قيمتها عما لو استخدم طريقة الدرجات

الخام، وعندما يكون عدد فئات كل من المتغيرين (١٠) تقل قيمة معامل الارتباط بقـــدر ٣%.

ويمكن تصحيح الأخطاء الناتجة عن تبويب البيانات لأي عدد من فئات كل من المـــتغيرين بقسمة معامل الارتباط الناتج من استخدام طريقة الترميز على مقدار ثابت يساوي عـــدد هذه الفئات، وقد أعد بعض علماء الإحصاء قائمة من الثوابت التي يمكن أن يستخدمها الباحث لإجراء تصحيح معامل الارتباط عندما تبوب البيانات في فئات مختلفـــة الـــسعة بالنسبة للمتغيرين (س)، (ص)، وهذه الثوابت بالجدول الآتي:

عدد الفئات	معامل التصحيح	مربع معامل التصحيح
٢	٠,٨١٦	٠,٦٦٧
٣	٠,٨٥٩	٠,٧٣٧
٤	٠,٩١٦	٠,٨٣٩
٥	٠,٩٤٣	٠,٨٩١
٦	٠,٩٦٠	٠,٩٢٣
٧	٠,٩٧٠	٠,٩٤١
٨	٠,٩٧٧	٠,٩٥٥
٩	٠,٩٨٢	٠,٩٦٤
١٠	٠,٩٨٥	٠,٩٧٠
١١	٠,٩٨٨	٠,٩٧٦
١٢	٠,٩٩٠	٠,٩٨٠
١٣	٠,٩٩١	٠,٩٨٣
١٤	٠,٩٩٢	٠,٩٨٥
١٥	٠,٩٩٤	٠,٩٨٧

فإذا افترضنا أننا حصلنا على معامل ارتباط = ٠,٦١ من بيانات مبوبة عدد فئات المـــتغير (س) فيها = ٨، وعدد فئات المتغير (ص)= ٩، فعندئذ يمكن الرجوع إلى هذا الجـــدول

لمعرفة قيمة كل من معاملي التصحيح في الحالتين وهما: ٠,٩٧٧، ٠,٩٨٢ على الترتيب. ولإجراء تصحيح معامل الارتباط الذي حصلنا عليه وهو ٠,٦١ نطبق الصيغة التالية:

$$ر م = \frac{ر}{(ح س)\ (ح ص)}$$

حيث $(ر_{ح})$ترمز إلى معامل الارتباط بعد تصحيحه

،(ر) إلى معامل ارتباط قبل التصحيح

،($ح_{س}$)، ($ح_{ص}$) إلى معامل تصحيح (س)، (ص). ويمكن الحصول عليهما من الجـــدول المذكور.

فإذا طبقنا صيغة المعادلة المذكورة على قيمة عامل الارتباط ٠,٦١ نجد أن:

$$ر_{ح} = \frac{٠,٦١}{(٠,٩٨٢)\ (٠,٩٧٧)} = ٠,٦٢٦$$

أي أن معامل الارتباط بعد تصحيحه من الأخطاء الناتجة عن التبويب = ٠,٦٢٦

وبالطبع إذا تساوي عدد فئات كل من المتغيرين يتساوي معامل تصحيح كـــل منـــهما، وتصبح صيغة التصحيح السابقة كالتالي:

$$ر_{ح} = \frac{ر}{ح^{٢}}$$

وهذا يعني أن المقام قد أصبح مساوياً مربع معامل التـــصحيح لأي مـــن (س) أو (ص)، ويفضل تطبيق هذه الصيغة عندما يكون عدد فئات كل من المتغيرين (س)، (ص) أقل من (١٠) فئات، وبخاصة إذا كان عدد الفئات (٨) أو أقل. ويفيد تطبيق هـــذه الـــصيغة في الحصول على قيمة أكثر دقة لمعامل الارتباط عندما تكون قيمته كبيرة، أما إذا كانـــت قيمته صغيرة، وبخاصة إذا كان حجم العينة المستخدمة صغيراً أيضاً فلن يفيـــد كـــثيراً تطبيق هذه الصيغة. ويجب أن يراعي الباحث أن معاملات التصحيح المبنيـــة بالجـــدول

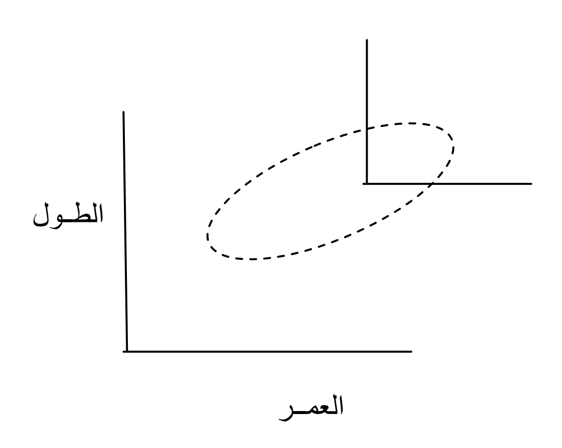
السابق قد أعدت بحيث تستخدم بوجه خاص في الحالات التي تكون فيها الفئات متساوية السعة ومنتصفات الفئات تمثل التكرارات، وأن يكون توزيع كل من المتغيرين اعتداليا.

العوامل التي تؤثر في معامل ارتباط بيرسون:

(١) إن إضافة أو طرح مقدار ثابت –لا يساوي صفراً –إلى أو من كل درجة من درجات أحد توزيعي المتغيرين أو كليهما، وكذلك الضرب في – أو القسمة على– مقدار ثابت لا يغير من قيمة معامل الارتباط. أي أن قيمته لا تتغير بتغير نقطة الأصل ووحدة ميزان القياس. والحقيقة أنه يمكن باستخدام هذه النتيجة في تبسيط العمليات الحسابية بأن نطرح مقداراً ثابتاً من كل درجة من درجات أحد المتغيرين أو كليهما إذا كانت قيم الدرجات كبيرة دون أن تتغير قيمة معامل الارتباط. كما أن هذه النتيجة تعني أنه يمكن إيجاد معامل الارتباط بين متغيرين مهما اختلفت وحدات قياس كل منهما، فقيمة معامل الارتباط بين العمر والطول لا تختلف سواء كانت وحدات العمر المستخدمة هي الأعوام أو الشهور، ووحدات الطول هي الأقدام أو السنتيمترات، وبالتالي فإن عدم تأثر معامل الارتباط بتغيير وحدة القياس أو نقطة الأصل لأي من المتغيرين أو كليهما يجعل معامل الارتباط من المقاييس الإحصائية ذات الأهمية التطبيقية الكبيرة.

(٢) تتأثر قيمة الارتباط بمدى تباين درجات كل من التوزيعين. فقيمة معامل الارتباط المحسوبة من مجموعة الدرجات المتباينة إلى حد كبير تكون أكبر من قيمته إذا كانت مجموعة الدرجات متقاربة في أحد المتغيرين أو كليهما،فمثلاً إذا حسبنا معامل الارتباط بين نسب ذكاء ودرجات تحصيل مجموعة من الطلاب ذوي الاختلاف الواضح في قدراتهم فإن قيمة معامل الارتباط قد تكون مرتفعة عما لو كانت مجموعة الطلاب من المتفوقين عقلياً، فمعامل الارتباط في هذه الحالة من المحتمل أن تكون قيمته منخفضة جداً بسبب تجانس المجموعة. وهذا يوضح أن قيمة معامل الارتباط بين متغيرين يكون لها معني فقط إذا حدد الباحث طبيعة وتكوين المجموعة موضع البحث.

وأحيانا يحصل الباحث على معامل ارتباط منخفض زائف أو وهمي Spurious Correlation ناتج عن تضييق مدى قيم أحد المتغيرين، فمثلاً إذا كان الباحث مهتماً بإيجاد العلاقة بين عمر وطول مجموعة من الأطفال الذين تتراوح أعمارهم بين ٣ أعوام، ١٦ عاماً، فإنه سيحصل بلا شك على معامل ارتباط مرتفع بين المتغيرين. أما إذا ضيق مدى أحد هذين المتغيرين بأن أوجد معامل الارتباط بين العمر والطول بالنسبة للأطفال الذين تتراوح أعمارهم بين ٩، ١٠ أعوام فقط، فإنه سيجد أن معامل الارتباط قد انخفض إلى حد كبير ويمكن توضيح ذلك بالشكل التالي:



فهذا الشكل الانتشاري يوضح قيمة مرتفعة لمعامل الارتباط بين العمر والطول على مدى متسع لكل منهما، ويوضح انخفاض قيمته عند تضييق مدى المتغيرين، ومن هذا الشكل نجد أن قيمة معامل الارتباط بين المتغيرين تكون كبيرة إذا أخذنا في الحسبان المدى الكلي لهما، أما إذا نظرنا إلى الجزء العلوي الأيمن من الشكل فسنجد أن هذه القيمة قد انخفضت بسبب تضييق هذا المدى. وكثيراً ما يواجه الباحث مثل هذه المشكلة وهي مشكلة تضييق أو بتر المدى الكلي لأحد المتغيرين أو كليهما، إذ إن كثيراً من الباحثين يجرون أبحاثهم على فئة معينة من الشباب (كشباب الجامعات مثلاً) ، وهؤلاء الشباب يكونون بمثابة مجموعة متجانسة من

حيث المتغيرات أو الخصائص التي تؤثر في الظاهرة محل القياس، وقد ينتج عن ذلك انخفاض قيمة معامل الارتباط بسبب تضييق المدى، وكلما كانت العينة التي أجريت عليها الدراسة أكثر تجانساً انخفضت قيمة معامل الارتباط (كأن يتم إجراء الدراسة على الطلاب المتفوقين مثلاً)، فهذه العينة متجانسة المفردات، وهـــذا التجـــانس في العينة يعني تضييق المدى

(٣) لكي تصل قيمة معامل ارتباط بيرسون إلى قيمتها العظمى (+١)، (−١) يجب أن يكون توزيعا المتغيرين له نفس الشكل. فمثلاً إذا كان أحد المـــتغيرين متـــصلاً والآخر من نوع المتغير الثنائي (أي الذي تكون قيمته إما واحداً صحيحاً أو صفراً على سبيل المثال)، فإن معامل الارتباط سوف يكون دائماً أقـــل مـــن الواحـــد الصحيح، وبالمثل إذا كان توزيع أحد المتغيرين ملتوياً إلى اليسار بينما كان توزيع المتغير الأخر ملتوياً إلى اليمين، فإن معامل الارتباط سوف يكون أيضا أقل مـــن الواحد الصحيح.

تفسير معامل ارتباط بيرسون:

يتضح مما سبق أن معامل الارتباط بين متغيرين هو قيمة مجردة تعبر عـــن العلاقـــة القائمة بين المتغيرين بحيث تنحصر بين (+١)، (−١)، ويعبر عن قيمـــة معامـــل الارتباط بكسر عشري.

وهنا يجب أن نحذر الباحث من الوقوع في خطأ تفسير معامل الارتباط على أنـــه قيمة مطلقة مثل القيمة للطول أو الوزن مثلاً، أو على أنه نسبة مئويـــة، فمـــثلاً معامل الارتباط (٠,٢٥) لا يعد نصف معامـــل الارتبـــاط (٠,٥٠)، ومعامـــل الارتباط (٠,٥٠) لا يعد نصف معامل الارتباط الذي قيمته واحد صحيح. كما أن الفرق بين معاملي الارتباط (٠,٤٠)، (٠,٦٠) لا يساوي الفرق بين معاملي الارتباط (٠,٧٠)، (٠,٩٠)، فمعامل الارتباط هو مقدار مجرد ولا يقاس علـــى ميزان خطي وحداته متساوية. كما لا يجب تفسير معامل الارتباط على أســـاس وحدات الدرجات الأصلية، حيث إن قيمة معامل الارتباط تكون مستقلة — كما

سبق أن ذكرنا – عن الوحدات التي يقاس بها المتغيران والقيم التي يأخذها كل منهما.

وأحيانا يعتبر الباحث أن معامل الارتباط تنحصر قيمته بين ٠,٣٠، ٠,٧٠ متوسط القيمة، أي يعبر عن علاقة ارتباطية متوسطة، بينما يعتبر أن معامل الارتباط الذي تقل قيمته عن ذلك منخفضاً، أما إذا زادت قيمته عن ذلك فإنه يعتبر مرتفعاً، لكن هذه الاعتبارات خاطئة من وجهة نظر الأساليب الاستدلالية في تحليل البيانات ، فدلالة معامل الارتباط هي دالة لحجم العينة، حيث إن قيمة معامل الارتباط المرتفعة التي يحصل عليها الباحث باستخدام عينات صغيرة ربما لا يكون لها أي معني على الإطلاق من ناحية الاستدلال على الارتباط في المجتمع الأصل الذي استمدت منه هذه العينات.

كما أن هذه الاعتبارات خاطئة أيضا من وجهة نظر الأساليب الوصفية في تحليل البيانات، حيث إن طبيعة كل من العينة والمتغيرات موضع البحث، والغرض من استخدام معامل الارتباط تعد من العوامل التي تحدد ما إذا كانت قيمة معامل الارتباط مرتفعة أو منخفضة. ويجب أن يلاحظ الباحث أيضا أن مقدار العلاقة بين متغيرين لا تعتمد على إشارة معامل الارتباط، فمعامل الارتباط (-٠,٧٠) يعبر عن نفس مقدار العلاقة بين متغيرين معامل الارتباط بينهما (+٠,٧٠) فالفرق بينها يكون في اتجاه العلاقة وليس في قيمة العلاقة.

وربما يواجه الباحث أيضاً مشكلة أخرى عند تفسير معامل الارتباط تنتج من فكرة إضافة أو طرح مقدار ثابت إلى أو من كل قيمة من قيم أحد المتغيرين لا تغير من قيمة معامل الارتباط. فإذا افترضنا أن الباحث أراد تحديد العلاقة بين درجات اختبار طبق على المجموعة نفسها في مرتين مختلفتين، فإذا حصل على معامل ارتباط مرتفع ربما تكون درجات المجموعة في المرة الثانية أعلى أو أقل من درجاتها في المرة الأولى، وبالمثل معامل الارتباط المرتفع بين درجات مجموعة من الأطفال في اختبار في القدرة على القراءة، واختبار في القدرة العددية ليس دليلاً على أن نمو القدرتين عندهما متكافئ، فمعامل الارتباط هو قيمة تدل على التغاير

أو التباين المتلازم Concomitant Variation بين المتغيرين، ولا يشير إلى مقـــدار المتغيرين.

ومن الطرق المفيدة في تفسير القيم المختلفة لمعامل الارتباط (ر) هو تربيـــع هـــذه القيم أي الحصول على قيمة ($ر^٢$) ، والمقدار ($ر^٢$) هو النسبة بين التباين الكلـــي لأحد المتغيرين والجزء من هذا التباين الذي يمكن التنبؤ به باستخدام المتغير الثاني. أي أن ($ر^٢$) هي الجزء من التباين أحد المتغيرين الذي يمكن أن نتنبأ به باســـتخدام المتغير الثاني. فإذا كان معامل الارتباط بين متغيرين هو(٠,٧٠٧) مثلاً فإن $ر^٢$ = $(٧٠٧)^٢$ =٠,٥ تقريبا، وعندما ر= ٠,٥٠ فإن $ر^٢$= ٠,٢٥، ولذلك فإنه يمكـــن اعتبار أن معامل الارتباط ٠,٧٠٧ ضعف معامـــل الارتبـــاط ٠,٥٠ حيـــث إن نسبة($ر^٢$) في الحالتين هي (١:٢) تقريبا.

خامساً: معامل الارتباط الجزئي

يمثل الارتباط الجزئي Partial Correlation أحد تطبيقات الضبط الإحصائي، وجوهر هذا الارتباط قياس العلاقة المستقيمة بين متغيرين بعد عزل تأثير المتغيرات الأخرى، لنفرض أن لدينا المتغيرات الثلاثة الآتية:

المتغير الأول: المعرفة بالأحداث السياسية

المتغير الثاني:معدل استخدام التليفزيون

المتغير الثالث: معدل قراءة الصحف

باستخدام معامل الارتباط الجزئي يمكننا حساب معامل الارتباط بين أي متغيرين اثـــنين من تلك المتغيرات بعد تثبيت أثر المتغير الثالث بحيث لا يؤثر في قيمة معامل الارتباط. على سبيل المثال يمكن معرفة الارتباط بين المعرفة بالأحداث الـــسياسية ومعـــدل اســـتخدام التليفزيون بعد تثبيت معدل قراءة الصحف، كما يمكن معرفة الارتبـــاط بـــين المعرفـــة بالأحداث السياسية و معدل قراءة الصحف بعد تثبيت معدل استخدام التليفزيـــون.... لقد تم عزل تأثير متغير ثالث من العلاقة بين متغيرين. إن معامل الارتباط الجزئي في هـــذه الحالة يكون من الرتبة الأولى First order partial correlation وذلك لأن هذا أبـــسط مستوى للارتباط الجزئي ، حيث المطلوب معرفة معامل الارتباط بين متغيرين بعد تثبيت تأثير متغير ثالث في تلك العلاقة، أي أن العمليات الإحصائية تتناول ثلاثة متغيرات، مـــع تثبيت أحدهم ثم بحث العلاقة بين الاثنين الآخرين.

لكن الكثير من الدراسات العلمية لا تقتصر على ثلاثة متغيرات، وإنما تشمل متغيرات عديدة ، (مثل: المعرفة السياسية، معدل مشاهدة التليفزيون، معدل قراءة الصحف، معدل سماع الراديو، معدل استخدام الإنترنت، مطالعة الكتب المعنية بالشئون العامة، الاتجاه السياسي، المشاركة في مؤسسات المجتمع المدني......الخ)، وكثيراً ما تقتضي الدراسة مثلاً رصد وتحليل العلاقة بين المعرفة السياسة ومعدل قراءة الصحف بعد عزل تأثير المتغيرات الأخرى (معدل مشاهدة التليفزيون، معدل سماع الراديو، معدل استخدام الإنترنت، مطالعة الكتب المعنية بالشئون العامة، الاتجاه السياسي، المشاركة في مؤسسات المجتمع المدني) في هذه الحالة يكون معامل الارتباط الجزئي من الرتبة الثانية Second order partial correlation

ومن الطبيعي أن يتم رصد معامل الارتباط (Correlation)بين المتغيرين المعنيين قبل تثبيت المتغير الثالث، وبموجب هذا التثبيت والحصول على قيمة معامل الارتباط الجزئي (Partial Correlation)، بعد التثبيت نكون أمام الاحتمالات الآتية :

- أن تكون قيمة معامل الارتباط الجزئي أقل من قيمة معامل الارتباط، وهذا يعنى أن المتغير المعزول كان يزيد العلاقة بين المتغيرين الآخرين (على سبيل المثال، فإن معامل ارتباط بيرسون بين المعرفة السياسية وقراءة الصحف كان ٠,٥٦ وبعد عزل تأثير مشاهدة التليفزيون تبين أن قيمة معامل الارتباط الجزئي بين المعرفة السياسية وقراءة الصحف أصبحت ٠,٤٣ لقد انخفض الارتباط بين المتغيرين ، هذا يعني أن مشاهدة التليفزيون كان لها تأثير موجب في العلاقة بين المعرفة السياسية وقراءة الصحف .

- أن تكون قيمة معامل الارتباط الجزئي أعلى من قيمة معامل الارتباط، وهذا يعنى أن المتغير المعزول كان يضعف العلاقة بين المتغيرين الآخرين (على سبيل المثال، نفرض أن معامل ارتباط بيرسون بين المعرفة السياسية وقراءة الصحف كان ٠,٥٦ وبعد عزل تأثير استخدام الإنترنت تبين أن قيمة معامل الارتباط الجزئي بين المعرفة السياسية وقراءة الصحف أصبحت ٠,٦٧ لقد ارتفع الارتباط بين المتغيرين بعد تثبيت استخدام الإنترنت، هذا يعني أن استخدام الإنترنت كان له تأثير سالب في العلاقة بين المعرفة السياسية وقراءة الصحف .

- أن تتساوي قيمة معامل الارتباط الجزئي مع قيمة معامل الارتباط، وهذا يعنى أن المتغير المعزول لا تأثير له في العلاقة بين المتغير المستقل والمتغير التابع .

وكثيراً ما يفسر الارتباط الجزئي باستخدام مفهوم التباين المشترك ، وقد سبقت الإشارة إلى أنه قبل تطبيق الارتباط الجزئي يكون لدينا معرفة بقيم معاملات الارتباط بين المتغيرات، ومن الضروري رصد معامل الارتباط بين المتغير الأول والمتغير الثاني قبل تثبيت

المتغير الثالث، لنفرض أن قيمة معامل ارتباط بيرسون بـــين هـــذين المـــتغيرين كانـــت =٠,٦٤٧١ وعند تطبيق الارتباط الجزئي –أي تثبيت المتغير الثالث وبحث الارتباط بـــين المتغير الأول والمتغير الثاني- تبين أن قيمة الارتباط الجزئي ٠,٦٨٩٦

إن التباين المشترك الناتج عن إسهام المتغير الثالث يمكن الحصول عليه بسهولة من خلال مربع قيمة الارتباط الجزئي مطروحاً منه مربع قيمة ارتباط بيرسون

أي (٠,٦٨٩٦)٢ – (٠,٦٤٧١)٢

أي ٠,٤٧٥٥ – ٠,٤١٨٧ = ٠,٠٦

بقسمة هذه القيمة (٠,٠٦) على مربع معامل ارتباط بيرســـون والـــضرب في ١٠٠ نحصل على النسبة المئوية للارتباط الناتج عن إسهام المتغير الثالث، وكما سبقت الإشـــارة فإن مربع معامل ارتباط بيرسون (أي قبل تثبيت المتغير الثالث) يبلغ ٠,٤١٨٧ وبالتـــالي فإن:

٠,٠٦÷ ٠,٤١٨٧ ×١٠٠ = ١٤,٣%

أي أن النسبة المئوية للارتباط الناتج عن إسهام المتغير الثالث هي ١٤,٣% وهذا يعـــني أن النسبة المئوية للارتباط الناتج عن إسهام متغيرات أخرى يساوي ٨٥,٧ %

ويختلف الارتباط الجزئي partial correlation عن الارتباط شبه الجزئي semi partial correlation ، أو ما يعرف أحياناً بارتبـــاط الجـــزء part correlation ، ذلـــك أن الارتباط شبه الجزئي أو ارتباط الجزء يعنى عزل تأثير متغير معين من (أحد) المتغيرين اللذين نبحث العلاقة بينهما وليس من هذين المتغيرين مجتمعين،ففي مثالنا السابق تم رصد معامل الارتباط الجزئي بين المعرفة بالأحداث السياسية ومعدل استخدام التليفزيون بعد تثبيـــت معدل استخدام الراديو، لقد تم تثبيت معدل استخدام الراديو بالنسبة للمتغيرين الآخـــرين (المعرفة بالأحداث السياسية ومعدل استخدام التلفزيون)، وهنا نكون قد طبقنا الارتبـــاط الجزئي، أما في الارتباط شبه الجزئي أو ارتباط الجزء، فإننا نقوم بتثبيت معدل استخدام الراديو عن معدل استخدام التليفزيون فقط، وليس عن المعرفة السياسية، إن ذلـــك يعـــنى إبقاء تباين متغير المعرفة السياسية كما هو ، وقد يرى الباحث أن مبررات ذلك تتمثل في وجود شكوك حول فاعلية الراديو في المعرفة السياسية. ويلعب الارتباط شبه الجزئي ، أو ارتباط الجزء دوراً هاماً الكثير من التحليلات الإحصائية الأخرى مثل الارتبـــاط المتعـــدد والتحليل العاملي،وعلى الرغم من أهمية الارتباط الجزئي وشبه الجزئي إلا أنه لا يعنى علاقة السبب بالنتيجة، فوجود ارتباط دال إحصائياً لا يعني أن أحد المتغيرين سبب أو نتيجـــة للمتغير الآخر.

الفصل الثالث
الانحــــــــــدار

تقديم:

يمكن الاستفادة من معرفة معامل الارتباط بين متغيرين في عملية التنبؤ prediction، وتتم هذه العملية من خلال الانحدار Regression، فإذا علمنا مثلاً عدد الساعات التي يقضيها المفحوصون يومياً في قراءة الصحف كذلك عدد الساعات التي يقضونها يومياً في مشاهدة التليفزيون، فإن بالإمكان إيجاد معامل الارتباط بين هذين المتغيرين (قراءة الصحف ومشاهدة التليفزيون)، فإذا كان معامل الارتباط عالياً أو جوهرياً، فإننا يمكننا التنبؤ بالوقت الذي يقضيه أي مفحوص في مشاهدة التليفزيون إذا علمنا الوقت الذي يقضيه في قراءة الصحف، كما يمكننا معرفة الوقت الذي يقضيه في قراءة الصحف بمعلومية الوقت الذي يقضيه في مشاهدة التليفزيون. ويغطي تحليل الانحدار مدى واسعاً من التطبيقات والاستخدامات بدرجة يصعب معها الإحاطة به في كتاب واحد، فما بالنا بفصل في كتاب!. من هنا سيتم التعريف بالجوانب الأساسية للانحدار من خلال نموذجين فقط، الأول هو الانحدار الخطي البسيط، والثاني هو الانحدار اللوجستي

أولاً: الانحدار الخطى البسيط

الانحدار الخطى البسيط هو أسلوب إحصائي يهدف إلى التنبؤ بقيمة متغير معين هو المتغير التابع (ص) بمعلومية متغير آخر هو المتغير المستقل (س) ، وبذلك فإن الانحدار الخطي البسيط – كما يفهم من اسمه– هو أبسط مستويات الانحدار،حيث يتضمن متغيرين فقط أحدهما هو المتغير المستقل والثاني هو المتغير التابع . ولإجراء الانحدار الخطى البسيط لا بد أن تكون البيانات كمية Quantitative سواء فيما يخص المتغير التابع أو فيما يخص المتغير المستقل (وهذا يعني أن الارتباط المناسب هو معامل ارتباط بيرسون) .مثال ذلك تطبيق الانحدار الخطى البسيط على الدخل الشهري من جهة، والمبلغ الذي يتم إنفاقه على شراء الصحف اليومية من جهة ثانية. إن الدخل المالي الشهري يعامل كمتغير مستقل independent أما المبلغ الذي يتم إنفاقه على شراء الصحف اليومية فيعامل كمتغير تابع dependent

من جهة أخرى فإن استخدام الانحدار الخطي البسيط يتطلب أولاً التأكد من وجود ارتباط جوهري (Significant Correlation) بين المتغير التابع والمتغير المستقل، فإذا لم

يوجد هذا الارتباط، يصبح استخدام الانحدار الخطي البسيط غير ذي معنى، ذلك أنــه إذا كان معامل الارتباط بين المتغيرين يساوى الصفر، فإن هذا يعني أنه لا يوجد علاقة بين المتغيرين ، أي أننا لا نستطيع التنبؤ بقيمة أحد المتغيرين اعتماداً على معرفتنا بقيمة المــتغير الآخر، أما إذا كان معامل الارتباط أكبر من الصفر (بمعنى أنه هذا الارتباط دال إحصائياً) فإن ذلك يعني وجود علاقة جوهرية، وبالتالي يمكن التنبؤ، وتزداد دقة التنبؤ وقيمة الانحدار كلما ارتفعت قيمة معامل الارتباط .

وقد يتم التعامل مع المتغير الأول على أنه متغير مستقل والمتغير الثاني على أنه متغير تابع، كما قد يتم العكس، بمعنى التعامل مع المتغير الأول على أنه متغير تابع والمتغير الثــاني على أنه متغير مستقل، الأمر الذي يؤكد ضرورة أن يكون لدى الباحث مبررات منطقية لاعتبار هذا المتغير أو ذاك متغيراً مستقلاً أو تابعاً، فقد يكون هناك ارتباط طردي موجب بين معدل استخدام وسائل الإعلام (المتغير س) من جهة والمعرفة السياسية(المتغير ص) من جهة ثانية، بمعنى أن المعرفة السياسية لدى المفحوصين تزداد كلما زاد استخدامهم لوسائل الإعلام، إن وجود ارتباط طردي موجب بين هذين المــتغيرين لا يعــني بالــضرورة أن أحدهما سبب في الآخر، أو أن أحدهما نتيجة للآخر، كل ما في الأمر أن هناك تلازماً بين المتغيرين، فإذا زاد استخدام وسائل الإعلام زادت المعرفة السياسية، وإذا زادت المعرفــة السياسية زاد استخدام وسائل الإعلام.

لنفرض أن هناك دراسة علمية أجريت على عينة عشوائية من المفحوصين، وتبين منها أن متوسط درجة العينة على مقياس استخدام وسائل الإعلام هو (١٠) ومتوسط درجتــهم على مقياس المعرفة السياسية هو (٨) ويمكن أن نرمز إلى استخدم وسائل الإعلام بــصفته المتغير المستقل (س)، أما المعرفة السياسية فهي المتغير التابع (ص) .فإذا أردنا التنبؤ بدرجة المعرفة السياسية (ص) من درجة استخدام وسائل الإعلام (س) سمى هذا النوع من التنبؤ بانحدار (ص) على (س)، أما إذا أردنا التنبؤ بدرجة استخدام وسائل الإعلام (س) مــن درجة المعرفة السياسية (ص) سمى هذا النوع انحدار (س) على (ص)، بمعنى أوضح، فــإن انحدار (ص) على (س) يعني التنبؤ بقيم (ص) من واقع معرفتنا بقيمة (س)، أما انحــدار

(س) على (ص) فيعني التنبؤ بقيم (س) من واقع معرفتنا بقيمة (ص). فيما يخص انحدار (ص) على (س)، فإن معادلة الانحدار تأخذ الصورة الآتية :

$$\text{ص} = \text{ر} \times \frac{\text{ع ص}}{\text{ع س}} (\text{س} - \text{م س}) + \text{م ص}$$

حيث إن :

ص = الدرجة المجهولة (درجة المعرفة السياسية والتي نريد معرفتها من خلال الدرجة (س) المعروفة لدينا، وهى درجة استخدام وسائل الإعلام .

س = الدرجة المعلومة (درجة استخدام وسائل الإعلام)

ر = معامل الارتباط بين الدرجة س والدرجة ص

ع س = الانحراف المعياري لمتوسط الدرجة س

ع ص = الانحراف المعياري لمتوسط الدرجة ص

م س = متوسط درجة الاختبار (س)

م ص = متوسط درجات الاختبار (ص)

لنفترض أن معدل استخدام وسائل الإعلام (س) والمعرفة السياسية (ص) وأن الدراسة الوصفية للعلاقة بين هذين المتغيرين أتاحت لنا المعطيات الإحصائية الآتية عن المتغيرين (س) ، (ص) :-

ر = ٠,٩٠

ع ص = ٢,١٦

ع س = ٧,٥٦

م س = ١٠

م ص = ٨

.بموجب هذه المعطيات يمكن حساب انحدار (ص) على (س) كالآتي:

ص = ٠,٩٠ × ٢,٦١ ÷٧,٥٦ (س−١٠) + ٨

= ٠,٣١ (س−١٠) + ٨

= ٠,٣١ س − ٣,١ + ٨

= ٠,٣١س + ٤,٩

إذن

ص = ٠,٣١ س +٤,٩

وهذه هي معادلة انحدار (ص) على (س) أو معادلة التنبؤ المطلوبة، فإذا كانت فإذا كانت قيمة (س) تساوى ٣ فإن قيمة ص

= ٠,٣١× ٣ + ٤,٩ = ٥,٨٣ تقريباً

ويمكن الاستفادة من ذلك في رصد أي درجة للمتغير (ص) من واقع معرفة أي درجة للمتغير (س) ، فإذا كانت قيمة (س) يساوي ٦ مثلاً ، فإن قيمة (ص) تساوي:

٠,٣١ × ٦+ ٤,٩ =٦,٧٦

فإذا كانت القيمة (س) ترمز إلى الدرجة على مقياس استخدام وسائل الإعلام، وإذا كانت القيمة (ص) ترمز إلى الدرجة في المعرفة السياسية، فإن الشخص الذي حصل على ٦ في الدرجة على مقياس استخدام وسائل الإعلام نتنبأ أنه يحصل على ٦,٧٦ في الدرجة على مقياس المعرفة السياسية

أما فيما يخص انحدار (س) على (ص)، فإنه كما سبقت الإشارة يعني التنبؤ بقيمة (س) بناء على معرفة قيمة (ص) ، ويمكن الحصول على معادلة انحدار (س) على (ص)، بالطريقة نفسها، حيث إن:

$$س = ر \times \frac{ع\,س}{ع\,ص} (ص - م\,ص) + م\,س$$

بالاستفادة من المعطيات السابقة، عن المتغيرين (س)، (ص) وبالتعويض في تلك المعادلة، نجد أن:

س = ٠,٩٠ × ٧,٥٦÷ ٢,٦١ (ص-٨) + ١٠

= ٢,٦١ (ص-٨) +١٠

=٢,٦١ ص-٢٠,٩ + ١٠

= ٢,٦١ ص -١٠,٩

أي أن قيمة (س) تساوي ٢,٦١ص — ١٠,٩ فإذا كانت قيمة ص = ٦ ، فإنه يمكـــن التنبؤ بقيمة(س) كالآتي:

س = ٢,٦١ × ٦-١٠,٩ = ٤,٧٦

فإذا كانت القيمة (س) ترمز إلى الدرجة على مقياس استخدام وسائل الإعلام، وإذا كانت القيمة (ص) ترمز إلى الدرجة في المعرفة السياسية، فإن الشخص الذي حـــصل علـــى ٦ درجات في المعرفة السياسية يحصل على ٤,٧٦ في الدرجة على مقياس استخدام وسائل الإعلام

وتبدو أهمية الانحدار في الدراسات الإعلامية والاجتماعية عموماً، إذ إن هناك العديد مـــن القضايا الحيوية التي يقوم التخطيط لها على التنبؤ الدقيق، فنحن نريـــد أن نتنبـــأ بعمـــق التفاعل أو التواصل الأسري في ضوء معدل استخدام الإنترنت أو التليفزيون أو غير ذلك من وسائل الإعلام ، كما أنه من الأمور الجوهرية للقائمين على حملات ترشيد الاستهلاك مثلاً التنبؤ بمعدل التعرض للحملة الإعلامية Media Campaign في ضوء معدل استخدام وسائل الإعلام، وقد نشط الباحثون في علم الاجتماع بالتنبؤ باتجاه الشباب نحو الجريمة في ضوء معدل التعرض لأفلام وماد العنف في التليفزيون أو السينما أو الإنترنت، ويـــستفيد التخطيط الاجتماعي من تلك المعطيات الإحصائية في توعية الأسرة بشأن ضوابط سليمة فيما يخص علاقة الأبناء بوسائل الإعلام. من جهة أخرى، فإن الدراسات الاجتماعية تهتم بالعوامل المؤثرة في الأسرة كوحدة اجتماعية، وكثيراً ما تنحو هذه الدراسات إلى استخدام أسلوب تحليل الانحدار في التنبؤ بالعوامل ذات التأثير في معـــدلات الـــزواج الطـــلاق والاستفادة من ذلك في التخطيط الاجتماعي

ثانياً:الانحدار اللوجستي

كثيراً ما نجد كلمة لوجستي أو لوجستية في اللغة العربية وإن كانت كلمة إنجليزيةLogistic أي أننا نستخدمها دون تعريب مثلها في ذلك مثل الكلمات الأجنبية التي نستخدمها كما هي في اللغة الأجنبية الأصلية، ومن تلك الكلمات كلمة تليفزيون، راديو، تليفون .. الخ، وتأتى كلمة لوجستي / لوجستية غالباً كصفة، كأن نقول الدعم اللوجستي أو المساعدات اللوجستية، وذلك لوصف دعم مساند يشكل أهمية حيوية للعمليات الأساسية، فإذا افترضنا أن العمليات الأساسية في المعركة أو الحرب هي تفعيل الأسلحة والمعدات العسكرية، فإن النقل والتعبئة والتموين وغير ذلك من مستلزمات مساندة أو داعمة يدخل في عداد الدعم اللوجستي.

أما في الإحصاء فإن كلمة لوجستي Logistic تستخدم نسبة إلى (Log) وهو مختصر اللوغاريتمات Logarithms . إن اللوغاريتم ببساطة هو " الأس" الذي يجب أن يرفع إليه الأساس ليعطى عدداً معيناً ، وكمثال على ذلك، فإن :

٧ ^٣^ تنطق هكذا (٧ أس ٣) وهى تساوى ٣٤٣ ، أي أن ٧ ^٣^ (٧ أس ٣) =٣٤٣ .بمعنى أن الرقم ٧ مضروباً في نفسه ثلاث مرات (٧×٧×٧) يساوي ٣٤٣ ، ويسمى العدد ٧ بالأساس، أما الرقم ٣ فيسمى "بالأس" أي أن العدد ٧ يجب أن يرفع إلى الأس ٣ ليعطى العدد ٣٤٣ ، هذا "الأس" الذي يجب أن يرفع إليه العدد ٧ ليعطى العدد ٣٤٣ هو المعروف باللوغاريتم. وعند تحويل ٧ ^٣^ (أو ٧ أس ٣) من الصورة الأسية إلى الصورة اللوغارتمية يصبح:

٧ لو ٣٤٣=٣

وتقرأ كالآتي: لوغاريتم ٣٤٣ للأساس ٧ يساوى ٣

وهناك اللوغاريتمات الطبيعية التي أساسها الرمز (e) وعلى الرغم من أنه لا توجد قيمة محددة له تماماً إلا أنه قيمته التقريبية ٢,٧١٨٢٨١٨٠٠٠ ، أي ٢,٧١٨ علماً بأن هناك اللوغاريتمات العشرية التي يتخذ فيها الرقم ١٠ كأساس لها كما هو معروف.

وإذا كان الانحدار اللوجستي Logistic Regression ينسب إلى المصطلح Logarithms واختصاره Log فإن هذا التحليل يستخدم عندما يكون المتغير التابع ثنائيـــاً متقطعـــاً (Binary dichotomous) .بمعنى أن هذا المتغير له حالتان فقط مثل :

- مشاهدة التليفزيون مقابل عدم مشاهدة التليفزيون
- قراءة الصحف مقابل عدم قراء الصحف
- سماع الإذاعة مقابل عدم سماع الإذاعة
- استخدام الإنترنت مقابل عدم استخدام الإنترنت

ففي الحالة الأولى (مشاهدة التليفزيون) فإن المتغير الثنـــائي هـــو (علاقـــة المـــشاهد بالتليفزيون) هل يشاهد أو لا يشاهد، وفى الحالة الثانية نجد أن المتغير الثنائي هو (قـــراءة الصحف)، بمعنى علاقة الجمهور بالصحيفة هل يقرأ الصحيفة أو لا يقرأ الـــصحيفة ... وهكذا فيما يخص الإذاعة (يسمع الإذاعة مقابل لا يسمع الإذاعة)، وكذلك فيما يخـــص الإنترنت (يستخدم الإنترنت مقابل لا يستخدم الإنترنت

المنطق نفسه فيما يخص أي متغير ثنائي (امتلاك جهاز تليفزيون مقابل عدم امتلاك جهاز تليفزيون، قبول الاشتراك في الصحيفة رفض الاشتراك في الصحيفة، المشاركة في برنـــامج إذاعي مقابل عدم المشاركة في برنامج إذاعي،معرفة معلومة معينة مقابل عدم معرفة تلك المعلومة، المشاركة في الانتخابات مقابل عدم المشاركة في الانتخابات ... الخ)

فإذا كان المتغير التابع Dependent ثنائياً متقطعاً، ونريد تطبيق تحليل الانحدار فإن نموذج الانحدار المناسب هو الانحدار اللوجستي، وهنا تجد الإشارة إلى أن المتغير المستمر (غـــير المتقطع) يمكن إعادة تكويده Recoding بحيث يصبح متغيراً ثنائياً متقطعاً ، مثال ذلك نفترض أن نتائج أحد البحوث أفادت أن مشاهدة التليفزيون من واقع عينة قوامها ٤٦٨ مفردة (ن= ٤٦٨) جاءت كالآتي:

مشاهدة التليفزيون	ك	%
ثلاث ساعات	٧٢	١٥,٤
من ساعتين لأقل من ٣ ساعات	١٠٣	٢٢
من ساعة لأقل من ساعتين	٣٥	٧,٥
أقل من ساعة	٤٥	٩,٦
لا يشاهد التليفزيون	٢١٣	٤٥,٥
المجموع	**٤٦٨**	**١٠٠**

في هذه الحالة يمكن إعادة تكويد استجابات المفحوصين بحيث تضم فئتين فقط :

- يشاهد: العدد ٢٥٥
- لا يشاهد: العدد ٢١٣

فالمتغير المتصل (المستمر)- وهو عدد ساعات مشاهدة التليفزيون يومياً- تم تحويله إلى متغير متقطع ثنائي (له حالتان فقط)، وبموجب هذا التحويل يمكن تطبيق نموذج الانحدار اللوجستي .

ومن الممكن أن يكون لنموذج الانحدار اللوجستي تطبيقات بالغة الأهمية في مجال الإعلام والاتصال بالجماهير، فنحن نريد أن نعرف مثلاً احتمال أن يعمل جهاز مونتاج Final cut بكفاءة لمدة خمس سنوات أو لن يعمل بكفاءة هذه المدة، كما نريد أن نعرف ما إذا كانت الجريدة ستحقق رقما معينا في المبيعات أو لن تحقق هذا الرقم ، أو أن الجمهور سيشاهد التليفزيون بنسبة ٣٠% على الأقل أم لا ......الخ، كما يمكن تطبيق الانحدار اللوجستي بكفاءة عالية لمعرفة تأثير وسائل الإعلام وغيرها من العوامل الأخرى في موضوع ما، كأن يتم رصد مجموعة من المتغيرات الإعلامية (كمتغيرات مستقلة) لمعرفة الوعي السياسي أو عدم السياسي وفق مؤشرات معينة، أي أن المعرفة باحتمال حدوث أي من هذه الأمور تكون بناءً على رصد مجموعة المتغيرات ذات التأثير في هذا الاحتمال، وعادة ما يكون احتمال الحدوث مكوداً بالرقم (١) أما احتمال عدم الحدوث فيتم تكويده بالرقم (صفر) فإذا افترضنا أن حدوث المشاهدة التليفزيونية تأخذ الرقم (١) فإن

عدم المشاهدة يأخذ الرقم (صفر)، وهذا التكويد Coding يتم أثناء إدخال البيانات باستخدام أحد البرامج الإحصائية الحاسوبية المعروفة

ويرتبط الانحدار اللوجستي ارتباطاً وثيقاً بالاحتمالات فإذا افترضنا أن احتمال قراءة صحيفة يساوى ٧٠% فإن هذا يعنى أن احتمال عدم قراءة تلك الصحيفة هو ٣٠% وهنا تكون أرجحية قراءة الصحيفة

= ٠,٧ ÷ ٠,٣ = ٢,٣٣

أما أرجحية عدم قراءة الصحيفة فهي

=٠,٣ ÷ ٠,٧ = ٤٣,

وإذا كان الانحدار الخطى Linear Regression يبحث معدل التغير في المتغير التابع Y عندما يتغير المتغير المستقل X فإن الانحدار اللوجستي Logistic Regression يبحث في معدل التغير في لوغاريتم الأرجحية Log odds لحدوث المتغير التابع Y عندما يتغير المتغير المستقل X

إن نسبة الأرجحية هي احتمال وقوع الحدث مقسوماً على احتمال عدم وقوعه، فإذا كان احتمال مشاهدة التلفاز ٠,٦٨ فإن احتمال عدم المشاهدة هو ٠,٣٢ وتكون نسبة الأرجحية هي ٠,٦٨ ÷ ٠,٣٢ = ٢,١٢٥

ويتم التعرف على دلالة قيمة تلك النسبة من خلال "اختبار والد" Wald test فإذا كانت قيمة هذا الاختبار معنوية أو دالة Significant، فإن هذا يعنى أن المتغير المستقل له تأثير جوهري في ترجيح احتمال وقوع أو عدم وقوع الحدث، قد تكون هذه الدلالة عند مستوى ٠,٠٥ ($p < 0.05$) أو عند مستوى ٠,٠١ ($p < 0.01$)

وكمثال تطبيقي نفرض أن المتغير التابع المطلوب قياسه (التنبؤ به) هو "التصويت في الانتخابات البرلمانية" وعند إدخال البيانات فإن هذا المتغير يعامل كمتغير تابع Dependent variable ومن الواضح أنه من النوع الثنائي Dichotomous فالمفحوص إما أنه أدلى بصوته أو أنه لم يدلى بصوته في الانتخابات. أما مجموعة المتغيرات المستقلة Independent أو المنبئة فهي :

- مشاهد التليفزيون (الدعاية الانتخابية)
- سماع الإذاعة (متابعة شئون الانتخابات)
- قراءة الصحف اليومية الحكومية
- قراءة الصحف الحزبية
- الانتماء لحزب سياسي

وإذا كان المتغير التابع (الإدلاء بالصوت في الانتخابات) يتخذ الرقم الكودي (١) بمعنى التصويت في الانتخابات، أو الرقم الكودي (صفر) بمعنى عدم التصويت في الانتخابات، فإن مجموعة المتغيرات المستقلة لها قيم كمية معينة تعكس استجابات المفحوصين حسبما أفادوا بها، ومع إدخال الرمزيين الكوديين للمتغير التابع (الرقم صفر أو الرقم ١)، و إدخال القيم الكمية للمتغيرات المستقلة فإن نموذج الانحدار اللوجستي يتعامل مع تلك المتغيرات المستقلة كمتغيرات منبئة Predicting Variables ويسفر تطبيق هذا النموذج معطيات إحصائية هامة مثل: سالب ضعف دالة الإمكان (2loglikelhood -) حيث توجد قيمة لهذه الدالة في كل خطوة وذلك لاختبار كفاءة النموذج، كما تتضمن المعطيات الإحصائية قيمة R2 (مربع معامل الارتباط) بطريقة Cox &snell وكذلك بطريقة Nagelkerke بالإضافة إلى مربع كاي (Chi square) ودلالته....ويستعان بهذه المعطيات الإحصائية في رصد وتفسير النتائج، كما يمكن للباحثين الاستفادة بها في إجراء تحليلات إحصائية أكثر تطوراً

والجدول الآتي يوضح بعض المعطيات الإحصائية لنموذج الانحدار اللوجستي المقدر باعتبار أن الموقف من الانتخابات (التصويت أو عدم التصويت) متغير تابع (Y) ، أما المتغيرات المستقلة فهي: مشاهد التليفزيون ، سماع الإذاعة، قراءة الصحف اليومية الحكومية، قراءة الصحف الحزبية، الانتماء لحزب سياسي:

النموذج المقدر

	B	S.E	Wald	df.	Sig.	Exp (B)
مشاهدة التليفزيون	٠,٤٧١	٠,١٢٧	١٣,٨٤	١	٠,٠٠٠	٠,٦٢٤
سماع الراديو	٠,٤٧٢	٠,٢٠٣	٥,٤٤	١	٠,٠٢	١,٦٠٤
قراءة الصحف الحكومية	٠,٤٩٢	٠,١٦١	٩,٤	١	٠,٠٠٢	١,٦٣٦
قراءة الصحف الحزبية	٠,٥٩٢	٠,١٢٢	١٨,٨	١	٠,٠٠٠	,٥٨٩
الانتماء لحزب سياسي	٠,١٠٥	٠,١٣٩	٠,٥٧٨	١	٠,٤٤٧	٠,٠٠٩
الثابت	**٠,٩٦٣**	**٠,٣٧٧**	**٦,٥**	**١**	**٠,٠١١**	**٢,٦٢**

يلاحظ من هذا الجدول أن جميع قيم B موجبة (العمود الأول)، وهذه القيم هي معاملات النموذج المقدر بوحدات Log odds أما العمود الثاني S.E فيبين الخطأ المعياري لكل معامل من المعاملات B ، ويتضمن العمود الثالث معطيات اختبار Wald وهو في الحقيقة اختبار (كا^2) المبنى على معامل Wald لمعرفة معنوية قيمة كل معامل، ومن الجدول أيضاً وفي ضوء قيمة المعنوية .Sig نتبين أن الانتماء الحزبي غير دال إحصائياً (Sig.>0.05) أما بقية المتغيرات فهي دالة إحصائياً (Sig.<0.05) أخيراً، فإن القيم الموجودة في العمود (Exp B) تعبر عن نسبة الأفضلية لكل معامل، وهذه النسبة هي المعامل مرفوعاً في الأس (e) وقد سبقت الإشارة إلى أن قيمة هذا الأس ٢,٧١٨ تقريبا (اللوغاريتم الطبيعي)

ويتم الحصول على معادلة الانحدار اللوجستي بإضافة الثابت Constant إلى القيم الواردة في العمود الأول (B)، أما ناتج هذه المعادلة فيرمز له بالحرف Z ، وحسب المعطيات الإحصائية بهذا الجدول تكون قيمة Z أو (ز) كالآتي:

ز= ٠,٩٦٣+٠,١٠٥ الانتماء الحزبي + ٠,٥٩٢ قراءة الصحف الحزبية + ٠,٤٩٢ قراءة الصحف الحكومية + ٠,٤٧٢ سماع الراديو + ٠,٤٧١ مشاهدة التليفزيون

أما الصياغة الإنجليزية لتلك المعادلة فهي:

Z=0.963+0.105 political party belonging+0.592 reading the political parties newspapers+0.492 reading the governmental newspapers+0.472 radio listening +0.472 television viewing

ويستفاد بقيمة ناتج هذه المعادلة أي قيمة (Z) وكذلك بقيمة الأس الطبيعي (e) في تقدير احتمال حدوث المتغير التابع ، وذلك بموجب المعادلة:

$$Pro. = \frac{1}{1 + e^{(-Z)}}$$

وقد سبقت الإشارة إلى أن Z ترمز إلى قيمة ناتج معادلة الانحدار اللوجستي، كما أن الحرف e يرمز إلى قيمة اللوغاريتم الطبيعي وقدرها ٢,٧١٨ كما ذكرنا من قبل.إن ناتج تطبيق هذه المعادلة يسمى الاحتمال المقـــدر (The estimated Probability) أي احتمال وقوع الحدث، فإذا بلغت قيمة هذا الاحتمال (المقدر) حوالي ٠,٧٠ مثلاً، فـــإن ذلك يعني أن أرجحية وقوع الحدث تبلغ ٠,٧٠ وبالتالي فإن أرجحية عـــدم وقوعـــه ٠,٣٠ فإذا كانت قيمة الاحتمال المقدر في مثالنا السابق هي ٠,٧٠ فإن ذلك يعـــني أن المتغيرات المستقلة تزيد أرجحية التصويت في الانتخابات بنسبة ٠,٧٠ وهذه المـــتغيرات هي: مشاهد التليفزيون ، سماع الإذاعة، قراءة الصحف اليومية الحكومية، قراءة الصحف الحزبية، الانتماء لحزب سياسي، وغني عن البيان أن كافة العمليات الحسابية يقـــوم بهـــا الحاسب الآلي

الفصل الرابع
المقارنة بين المتوسطات

تقديم:

يعتمد الكثير من البحوث العلمية على المقارنات المختلفة، بل إن هناك أسلوباً علمياً في البحث تحت هذا المسمى (الأسلوب المقارن) ، وتأتي المقارنة بين المتوسطات باعتبارها من أبرز تطبيقات هذا الأسلوب، والمتوسطات هنا تعني المتوسط الحسابي Mean باعتباره أحد مقاييس النزعة المركزية، وهو أحد أبرز الأساليب الإحصائية المستخدمة في العلوم الاجتماعية عموماً بما في ذلك بحوث الاتصال، وكثيراً ما يستخدم المتوسط ضمن الإحصاء الوصفي لمعرفة متوسط القيمة الكمية لمجمل استجابات العينة على مقياس أو اختبار معين... في هذا الفصل نبذة تطبيقية للمقارنة بين المتوسطات من خلال التعريف بنماذج تطبيقية للمتوسط الحسابي، ثم المقارنة بين متوسطين (اختبار "ت" ، وأخيراً تحليل التباين أحادي الاتجاه.

أولاً: المتوسط الحسابي

يعتبر المتوسط الحسابي (Mean) من مقاييس النزعة المركزية التي يشيع استخدامها في بحوث الاتصال، كما أن استخدامه أمر أساسي ونحن بصدد الاستدلال الإحصائي من العينة عن المجتمع. وعلى الرغم من أن المتوسط الحسابي يتأثر بالقيم المتطرفة – بدرجة أكبر من مقاييس النزعة المركزية الأخرى – إلا أن هذا التأثر يكون في العينات الصغيرة، أي عندما تقل العينة عن ٣٠ مفردة (n< 30) في الوقت الذي نجد بحوث الاتصال تميل غالبا إلى أن تستمد بياناها من عينة كبيرة نسبياً (يعتمد أغلب تلك البحوث على عينات قوامها ٤٠٠ مفردة) .

ومن المعروف أن المتوسط الحسابي هو مجموع القيم الكمية مقسوماً على عددها، إنه ناتج قسمة المجموع الجبري لهذه القيم، فإذا كان لدينا ثلاثة أعداد هي : ٢٦،٥٤،٣٨ فإن المتوسط الحسابي يساوى ٣٨+٥٤+٢٦÷٣=٣٩,٣ . وإذا كان لدينا عينة من ٥٠٠ شخص وطبق عليهم استبيان يتكون من ٢٠ بنداً ، لكل بند ثلاث استجابات تتخذ الأرقام : ١، ٢، ٣ فإن ذلك يعني أن كل شخص سيحصل على درجة تتراوح ما بين ٢٠ إلى ٦٠ ، وبالتالي يكون لدينا عدد من الدرجات مساوياً لعدد مفردات العينة، أي أن لدينا ٥٠٠ درجة، فإذا جمعنا درجات جميع المفحوصين وقسمناها على ٥٠٠ يكون

الناتج هو المتوسط الحسابي، أي متوسط درجة استجابة العينة على الاستبيان. وفي بحوث الاتصال يستخدم المتوسط الحسابي على نطاق واسع لرصد وتحليل شدة وطبيعة استجابات المفحوصين على المقاييس والاختبارات والاستبانات، والمتغيرات التجميعية، فقد يكـــون هناك مقياس يقيس علاقة المشاهد بالتليفزيون أو اختبار يقيس المعرفة الـــسياسية لـــدى مشاهدي التليفزيون أو مستمعي الراديو أو قراء الصحف ... ويتضمن المقياس أو الاختبار مجموعة من البنود التي لها استجابات تأخذ قيماً كميـــة معينـــة Quantitative values وهناك أيضاً متغيرات تجميعية بمعنى تلك المتغيرات التي يتم إعدادها بحيث تشمل مـــتغيرات فرعية كثيرة يتم دمجها في متغير واحد هو المتغير التجميعي، أبرز مثال علـــى ذلـــك في بحوث الاتصال هو متغير (المستوى الاقتصادي الاجتماعي) وكذلك مـــتغير (العلاقـــة بوسائل الإعلام) ....، فمتغير المستوى الاقتصادي الاجتماعي مثلاً يتـــضمن مؤشـــرات عديدة مثل الدخل، التعليم ، منطقة السكن ، الاشتراك في نادى ثقافي ، امتلاك الأجهـــزة الحديثة ... الخ...إن هذه المتغيرات الفرعية يكون لكل منها قـــيم عدديـــة، بموجـــب استجابات المفحوص على تلك المتغيرات تتحدد القيم العددية أو الكمية الـــتي تعكـــس مستواه الاقتصادي الاجتماعي، وعندما يتم تجميع هذه القيم (العددية) لكل مـــتغيرات المستوى الاقتصادي الاجتماعي يمكن تصنيف المفحوص حسب مجمل تلك المتغيرات التي تجسد مؤشرات هذا المستوى، بموجب ذلك يمكن معرفة متوسط المـــستوى الاقتـــصادي الاجتماعي للعينة ككل ، وكذلك معرفة متوسط المـــستوى الاقتـــصادي الاجتمـــاعي للمجموعات الفرعية التي تضمها العينة (كمعرفة متوسط المفحوصين الذكور مقارنـــة بمتوسط الإناث، أو معرفة متوسط المفحوصين كثيفي المشاهدة على مؤشرات المـــستوى الاقتصادي الاجتماعي مقارنة بمتوسط المفحوصين منخفضي المشاهدة....)

في الوقت نفسه كثيراً ما يتطلب البحث مثلاً معرفة معنوية الفروق بين المفحوصين مـــن حيث استخدام وسائل الإعلام حسب المستوى الاقتصادي الاجتماعي (ككـــل ولـــيس حسب كل مؤشر على حدة)، وفي هذا الحالة يتم تـــصنيف العينـــة حـــسب المـــستوى الاقتصادي الاجتماعي (منخفض، متوسط، مرتفع)، ويتم استخراج المتوســـط الحـــسابي لاستجابات مجموعات العينة على مقياس استخدام وسائل الإعلام، ومن ثم يمكن المقارنة

بين تلك المجموعات ومعرفة ما إذا كانت بينها فروق جوهرية في استخدام وسائل الإعلام حسب المستوى الاقتصادي الاجتماعي، كما أن بحوث الإعلام كثيراً ما تستخدم مقاييس تتضمن بنوداً متعددة لقياس موضوع معين وبالتالي تبدو ضرورة رصد وتحليل متوسط مجمل استجابات العينة على تلك البنود مجتمعة وليس على كل بند بمفرده، فليس من المنطقي مثلاً تحليل استجابات العينة على كل سؤال أو بند بمفرده، لكن من المنطقي تحليل الاستجابات على مجموعة من البنود أو على مجمل البنود.

مثال تطبيقي:

لدينا عينة قوامها ٥٠٠ مفردة وأجابوا عل مقياس يتضمن عشرين برنامجاً تليفزيونياً، كان المطلوب من كل مفحوص أن يختار إجابة واحدة من الإجابات الثلاث الآتية بشأن كل برنامج:

- يشاهد بصفة منتظمة
- يشاهد بصفة غير منتظمة
- لا يشاهد

وعند تكويد أو ترميز تلك الاستجابات، يراعي أن تتخذ هذه الاستجابات الثلاث أرقاماً متدرجة تنازلية (٣ ، ٢ ، ١) على التوالي، وليس (١، ٢، ٣) المنطق في ذلك أن المشاهدة بصفة منتظمة تكون أكبر من المشاهدة بصفة عير منتظمة، كما أن المشاهدة غير المنتظمة تفوق عدم المشاهدة، وبالتالي لا بد أن تعكس الأرقام هذه الحقيقة، بمعنى أن الرقم الدال على المشاهدة المنتظمة يكون أكبر من الرقم الدال على المشاهدة عير المنتظمة، كما أن الرقم الدال على المشاهدة غير المنتظمة لا بد أن يكون أكبر من الرقم الدال على عدم المشاهدة، وبدون مراعاة ذلك، فإن النتائج تأتي مضللة تماماً (وهذا الخطأ يقع فيه الكثير من الباحثين)

إن الأرقام (٣،٢، ١) هي التي يتعامل معها الإحصاء ، وبناء على تلك الأرقام فإن الدرجة التي تعكس استجابات المفحوصين على المقياس المذكور تتراوح ما بين (٢٠) إلى (٦٠) ، أي أن كل مفحوص سيحصل على درجة لا تقل عن ٢٠ ولا تزيد عن ٦٠ — فلنتصور أننا قمنا بجمع درجات جميع المفحوصين ثم قسمنا هذا المجموع على عدد هؤلاء

المفحوصين (٥٠٠ مفحوص)، إن ناتج هذه القسمة هو المتوسط الحسابي لدرجة المفحوصين على المقياس المستخدم.

وتتيح الحزمة الإحصائية للعلوم الاجتماعية SPSS هذه العملية بسهولة وذلك من خلال إجراءات معينة تتضمن جمع درجات جميع المفحوصين آلياً ثم إعطاء الأمر المطلوب فنحصل على المتوسط الحسابي ، والانحراف المعياري، والمدى (أقل درجة وأعلى درجة)، ومن واقع تنفيذ هذه الإجراءات على بيانات أحد البحوث حصلنا على المعطيات الآتية :

م (المتوسط الحسابي) = ٣٦,٣٢

ع (الانحراف المعياري) = ٩,٧٢

مد (المدى) ٢١ إلى ٥٧، بمعنى أن أقل درجة كانت ٢١ بينما أعلى درجة كانت ٥٧

هكذا يمكن القول إن متوسط درجة العينة (ن=٥٠٠) على مقياس مشاهدة البرامج التليفزيونية هو ٣٦,٣٢ درجة بانحراف معياري ٩,٧٢ وقد تراوحت درجات المفحوصين ما بين ٢١ إلى ٥٧ . ويتضح من ذلك أن متوسط درجة المفحوصين يعادل حوالي ٥٤% من الدرجة العظمى للمقياس والتي تبلغ ٦٠ درجة كما سبقت الإشارة

وبالمثل ومن خلال الحزمة الإحصائية للعلوم الاجتماعية SPSS يمكن الحصول على متوسطات درجات مجموعات العينة كل مجموعة على حدة، وفي تجربة تطبيقية- على سبيل المثال- كانت العينة ٥٠٠ مفردة، وتتوزع حسب الجنس إلى ٢٤٨ ذكور مقابل ٢٥٢ وبحساب متوسط درجة مجموعة الذكور (ن = ٢٤٨) - من خلال اختيار مفردات تلك المجموعة - تبين أنه ٣٥,٥٧٦ درجة بانحراف معياري ٨,٤ . المنطق نفسه على مستوى مجموعة الإناث (ن = ٢٥٢) وقد تبين أن متوسط درجة تلك المجموعة ٢٩,١٩ بانحراف معياري ٥,٩٧

هكذا يمكن بسهولة الوقوف على صورة واضحة عن علاقة المشاهدين عينة البحث ببرامج التليفزيون، فعلى مستوى العينة ككل كان متوسط درجة المشاهدة ٣٢,٣٦ أما على مستوى مجموعة الذكور فقط فقد بلغ هذا المتوسط ٣٥,٥٧٦وعلى مستوى مجموعة الإناث كان هذا المتوسط ٢٩,١٩ وبالطريقة نفسها يمكن الحصول على متوسط درجة

جميع المجموعات الفرعية التي تضمها العينة أياً كان المتغير أساس التـصنيف: الجـنس (ذكور & إناث) محل الإقامة (ريف& حضر) السن (أقل من ٣٠ سنة & ٣٠ سـنة فأكثر) ... الخ

ويقترن المتوسط الحسابي عادة بـالانحراف المعيـاري Standard Deviation ويعـنى الانحراف عن المتوسط، وهو من أهم مقاييس التشتت، ففي المثال السابق الإشارة إليه إذا كان متوسط درجة مشاهدة العينة للبرامج التليفزيونية هو ٣٦,٣٢ فإن هناك مفحوصين حصلوا على درجة تقل عن هذا المتوسط ، وهناك أيضاً مفحوصون آخرون حصلوا على درجة تزيد عن هذا المتوسط،هذه الفروق (الانحراف عن المتوسط العام للعينة) تعبر عن الانحراف المعياري، ومن الناحية الإحصائية فإن مجموع الانحرافات عن المتوسط تساوى الصفر، فلنتأمل المثال الآتي :

لدينا درجات خمسة طلاب ، هذه الدرجات هي :

٢ ، ٣ ، ٤ ، ٥ ، ٦

متوسط هذه الـدرجات = ٢+٣+٤+٥+٦÷ ٥=٤ أي أن المتوسـط الحـسابي لهـذه الدرجات هو (٤) وعند حساب انحراف كل درجة عن المتوسط نجد أن:

- انحراف الدرجة الأولى = ٢-٤=-٢
- انحراف الدرجة الثانية = ٣-٤=-١
- انحراف الدرجة الثالثة = ٤-٤=صفر
- انحراف الدرجة الرابعة ٥-٤-= ١
- انحراف الدرجة الخامسة ٦-٤= ٢

فإذا حسبنا مجمل ناتج تلك الانحرافات نجد أنه يساوي الصفر:

-٢-١+ صفر +١+٢ = صفر

من هنا جاءت رؤية عالم الإحصاء (كارل بيرسون) بتربيع الفروق عن المتوسط وذلـك للتخلص من العلامات السالبة، وبتربيع الفروق عن المتوسط نجد أن القيم السابقة تصبح:

- مربع فروق الدرجة الأولى= -٢×-٢ = ٤
- مربع فروق الدرجة الثانية= -١×-١= ١

- مربع فروق الدرجة الثالثة= صفر× صفر= صفر
- مربع فروق الدرجة الرابعة= ١ × ١ = ١
- مربع فروق الدرجة الخامسة= ٢×٢ = ٤

وبجمع الناتج نجد أنه يساوي ٤+١+صفر +١+٤=١٠

ويكون متوسط هذا الناتج هو ١٠ ÷ ٥ = ٢

ويكون الانحراف المعياري هو جذر تلك القيمة أي

$\sqrt{٢}$ = ١,٤١٤

أي أنه من الناحية الإحصائية فإن الانحراف المعياري هو الجذر التربيعي لمتوسط مربعات الانحرافات عن قيمة المتوسط الحسابي

وهناك طرق إحصائية لحساب الانحراف المعياري للدرجات الخام وللدرجات التكرارية ، ولفئات الدرجات ، .. الخ، إن معرفة هذه الطرق (الإحصائية)في حساب الانحراف المعياري تحقق الفهم لما يعنيه، خاصة وأن البرامج الإحصائية الجاهزة تتيح حساب الانحراف المعياري بسهولة وسرعة،وعلى الرغم من أن الانحراف المعياري يتأثر بالدرجات المتطرفة (لاعتماده على مربع فروق هذه الدرجات عن المتوسط) إلا أنه من أهم مقاييس التشتت كما أنه يرتبط ارتباطاً وثيقاً بمقاييس إحصائية هامة مثل معامل الالتواء والارتباط والدرجات المعيارية والدلالة الإحصائية .... الخ .

ثانياً:المقارنة بين متوسطين (اختبار" ت")

تتم المقارنة بين متوسطينTwo Means من خلال اختبار "ت" أو ما يعرف بمعامل t test – فهو من أبرز المعالجات الإحصائية لإجراء تلك المقارنات التي تعتمد على المتوسط الحسابي .

وقد يستخدم t. test على مستوى العينة الواحدة (لمعرفة معنوية الفروق بين متوسط درجة العينة على مقياس معين وبين متوسط درجة العينة نفسها على مقياس آخر ، أو لمقارنة متوسط درجة العينة بمتوسط درجة المجتمع أو بمتوسط آخر مفترض، أو لمقارنة

متوسط درجة العينة على مقياس معين قبل وبعد إدخال مؤثر معين، أو بعد فاصل زمـــني بين تطبيق وآخر .

ففي المثال السابق الإشارة إليه كان متوسط درجة العينة (ن=٥٠٠) على مقياس مشاهدة البرامج التليفزيونية هو ٣٢,٣٦ وقد يريد الباحث أن يعرف ما إذا كان هذا المتوســط يختلف اختلافاً جوهرياً عن قيمة مفترضة قدرها ٣١,٢٤ بدرجة ثقة ٩٥%

من واقع تطبيق t test تبين أن قيمة t تبلغ ٣,١٦٢ .بمستوى معنوية يـــساوي ٠,٠٠٢ ودرجة حرية ٤٩٩ (t.=3.162; df= 499; p=0.002)، وتدل قيمة المعنوية (٠,٠٠٢) على أن هناك فروقاً جوهرية بين المتوسطين، أي أن الفروق بينهما حقيقية ولا ترجع إلى الصدفة

كما يستخدم t test لاختبار معنوية الفروق بين متوسط درجـــات عينـــتين (ن١) و(ن٢) Two sample t test كالمقارنة بين متوسط درجة الذكور ومتوسط درجـــة الإناث من حيث كثافة مشاهدة التليفزيون أو قراءة الصحف أو سماع الإذاعـــة ...الخ ، مثال ذلك أن يريد الباحث معرفة الفروق بين متوسط درجة الذكور ومتوسط درجـــة الإناث على مقياس مشاهدة برامج التليفزيون في المثال السابق ذكره ،وفى هذه الحالة فإن الحزمة الإحصائية تتيح البيانات الآتية :

مجموعة الذكور	مجموعة الإناث
ن = ٢٤٨	ن = ٢٥٢
م = ٣٥,٦	م = ٢٩,٢
ع = ٨,٤	ع = ٥,٩٧

وبتطبيق t test لعينين مستقلتين independent Sample t. test تبين أن قيمـــة ت = ٩,٨٣٥ .بمستوى معنوية يساوي صفر(p=.000) ودرجة حرية٤٩٨ أي أن هناك فروقـــاً جوهرية بين الجنسين من حيث كثافة مشاهدة البرامج التليفزيونية حيث تـــزداد كثافـــة المشاهدة بفروق جوهرية لدى الذكور (م=٣٥,٦) مقارنة بالإناث (م=٢٩,٢)

– وقد يستخدم t test لمعرفة الفروق بين متوسطين مقتـــرنين Paired Sample t test مثال ذلك حساب معنوية الفروق بين متوسط درجة عينه على مقياس مشاهدة التليفزيون ومتوسط درجة العينة نفسها على المقياس نفسه بعد مضى فترة زمنية، ففي دراسة عـــن معدل استخدام وسائل الإعلام كانت عينة الثبات ٣٠ مفردة (ن=٣٠) وقـــد تم تطبيـــق المقياس على تلك العينة مرتين بفاصل زمني قدره ثلاثة أسابيع بين التطبيق الأول والتطبيق الثاني ، وبمعالجة البيانات في التطبيقيين تم الحصول على المعطيات الآتية :–

التطبيق الأول	التطبيق الثاني
م = ١١٢,٤	م = ١١١,٣
ع = ٩,٨	ع = ١١,٥

علماً بأن العينة – كما سبقت الإشارة – كانت ٣٠ مفردة (ن= ٣٠)، كما أن درجة الحرية في هذه الحالة تساوي (ن– ١) أي ٣٠ – ١= ٢٩ وبحساب معنوية الفروق بين المتوسطين (متوسط درجة العينة في التطبيق الأول ومتوسط درجة العينة نفسها في التطبيق الثاني) تبين أن قيمة t = ٠,٧٤٥ ومستوى المعنوية Sig. = ٠,٤٦٢ ومن الواضـــح أن قيمة مستوى المعنوية تزيد عن ٠,٠٥ أي أنه لا توجد فروق جوهرية بين متوسط درجة العينة في التطبيق الأول ومتوسط درجة العينة على المقياس نفسه في التطبيق الثاني، وهـــذا يعني ثبات المقياس ، إذ إن استجابات العينة على المقياس ظلت كما هي تقريباً في التطبيق الثاني كما كانت في التطبيق الأول .

وفى دراسة أخرى كان الهدف رصد وتحليل اتجاهات الطـــلاب نحـــو مهنـــة الخدمـــة الاجتماعية وتم اختيار عينة قوامها ستون طالباً وطالبة تتوزع بالتساوي بين الجنسين، وتم قياس اتجاهات العينة قبل برنامج تدريب ميداني ومشاهدة عدد من الأفلام الوثائقية الـــتي تتناول دور الاختصاصين الاجتماعيين في المستشفيات والمدارس ومؤسسات رعاية الأيتام. في القياس الأول كانت شدة الاتجاه الإيجابي نحو مهنة الخدمة الاجتماعية قـــد جـــاءت بمتوسط قدرة ٦٨,٨٣ بانحراف معياري ٢١,١ أما في القياس الثاني– أي بعد التدريب الميداني ومشاهدة العينة لمجموعة الأفلام الوثائقية – فقد جاءت شدة الاتجاه نحـــو مهنتـــه الخدمة الاجتماعية بمتوسط قدرة ٨٤ وانحراف معياري ١٩,٤ ، ويلاحظ ارتفاع قيمة

المتوسط في القياس الثاني، الأمر الذي قد يعني أن الاتجاه الإيجابي نحو مهنة الخدمة الاجتماعية قد ارتفع بعد التدريب الميداني والتعرض لمواد وثائقية عن دور الاختصاصيين الاجتماعيين في مواقع مختلفة، لكن لا يمكن الجزم بذلك إلا بعد التأكد من أن الفروق بين المتوسطين جوهرية ولا ترجع إلى الصدفة، بمعنى أوضح ،فإن التساؤل هو : هل توجد فروق جوهرية بين متوسط الاتجاه نحو مهنته الخدمة الاجتماعية قبل وبعد التدريب والتعرض للأفلام الوثائقية؟ إن تطبيق t test بما يتيحه معطيات إحصائية يمكننا من الإجابة على هذا التساؤل، علماً بأن (ن = ٦٠ ، درجة الحرية = ٥٩)

وبتطبيق Paired Samples t. test جاءت قيمة t ٦,٨٥١ وهذه القيمة دالة إحصائياً بمستوى معنوية صفر (p=0.000)بما يشير إلى أن الفروق جوهرية ولا ترجع إلى الصدفة، وهذا يعنى أن الاتجاه الإيجابي نحو مهنة الخدمة الاجتماعية قد ارتفع ارتفاعاً جوهرياً بعد التدريب الميداني ومشاهدة الأفلام الوثائقية إذ أصبح ٨٤ بعد أن كان ٦٨,٨٣ قبل مشاهدة تلك الأفلام

في هذا الصدد يستخدم بعض الباحثين ما يعرف بنسبة الكسب المعدل (Modified gaining) لضمان الحصول على تقدير دقيق لفاعلية المؤثر (وهو في مثالنا المذكور التدريب المبرمج ومشاهدة الأفلام الوثائقية التي تتناول دور الاختصاصين الاجتماعيين في المستشفيات والمدارس ومؤسسات رعاية الأيتام)، ويتم الحصول على نسبة الكسب المعدل من خلال تطبيق معادلة بلاك (Blake) ، وهي :

$$\frac{ص - س}{د - س} + \frac{ص - س}{د}$$

حيث:

ص :تعني درجة الاختبار البعدي

س : تعني درجة الاختبار القبلي

د : تعني النهاية العظمى لدرجة الاختبار

فإذا كانت الدرجة العظمى لمقياس الاتجاه نحو مهنة الخدمة الاجتماعية هي ١٠٠ درجة، فإن نسبة الكسب المعدل تكون:

$$\frac{٨٤ - ٦٨,٨٣}{١٠٠ - ٦٨,٨٣} + \frac{٨٤ - ٦٨,٨٣}{١٠٠} = ٠,٦٤$$

وإذا كان اختبار t test من أكثر الاختبارات الإحصائية شيوعاً في الدراسات الاجتماعية والإنسانية، فإن هناك شروطاً لاستخدامه، وعلى الرغم من التأكيد على ضرورة الالتـــزام بتلك الشروط في بعض المراجع المتخصصة إلا أن البعض الآخر يرى أنها ليست صارمة أو جامدة، وأنه بالإمكان استخدام t. test طالما تم التأكد من جوانب أساسية أهمها ألا يقل عدد مفردات أي مجموعة أو عينة عن ٣٠ مفـــردة، وأن يكـــون حجـــم العينـــتين أو المجموعتين متقارباً (فلا يكون حجم العينة الأولى ٥٠٠ مفردة بينما حجم العينة الثانيـــة عشر مفردات مثلاً)، إذ إن التفاوت الكبير في حجم العينتين أو المجموعتين من شأنه أن يفضى إلى نتائج مضللة عند استخدام t test ويرجع ذلك إلى أن دلالة قيمة T تتأثر بحجم العينة (لأن درجة الحرية تعتمد على عدد مفردات العينة)، من جهـــة أخـــرى، فـــإن استخدام t test لمعرفة معنوية الفروق بين متوسطين يتطلب التأكد من تجانس العينيتين .بمعنى ألا توجد فروق جوهرية بين تباين (Variance) المجموعة الأولى وتبـــاين المجموعـــة الثانية. إن التعرف على ما إذا كانت هناك فروق جوهرية بين تباين المجموعة الأولى وتباين المجموعة الثانية يتم من خلال اختبار تجانس التبـــاين (Homogeneity of variance) وتعبر عنه دلالة إحصاءات ليفن Levine statistics فإذا كانت معنوية هذه الإحصاءات أقل من ٠,٠٥ فإن ذلك دليل قاطع على أن تباين المجموعتين يختلف اختلافـــاً جوهريـــا، وبالتالي تكون معطيات t test غير سليمة ، أما إذا كانت المعنوية أكبر من ٠,٠٥ فهـــذا يدل تقارب تباين المجموعتين أو العينتين. أما الشرط الآخر لاستخدام t test فهو التأكد أولاً من اعتدالية التوزيع في كلتا المجموعتين ، ويمكن التعرف على ذلك بالعديـــد مـــن الطرق منها قياس الالتواء Skewness لكل مجموعة فإذا كانت الالتواء أقل من الواحـــد الصحيح سواء بالسالب أو بالموجب – فإن ذلك يعنى أن التوزيع معتدلاً أو قريباً مـــن

الاعتدال، وبالتالي يمكن استخدامt test أما إذا كانت قيمة الالتـــواء تتجـــاوز الواحـــد الصحيح فإن ذلك يجعل معطيات t test غير دقيقة .

ثالثاً: تحليل التباين أحادي الاتجاه

يعتبر تحليل التباين (Variance Analysis) من أكثر الأساليب الإحصائية اســـتخداماً في بحوث العلوم الاجتماعية والإنسانية، ويعتمد على متوسطات قـــيم الاســـتجابة علـــى المقاييس والاختبارات المستخدمة كأدوات لجمع البيانات، ولتحقيق فهـــم أفـــضل لهـــذا الأسلوب لابد من الإشارة إلى ما يعنيه التباين Variance ، ومن ثم توضيح تحليل التباين.

تعريف التباين :

من الجدير بالتأكيد عليه هنا أن الحديث عن التباين في هذه الجزئية يعنى بنوع محدد مـــن التباين ألا وهو تباين الدرجات حول متوسط درجة العينة أو المجموعة، ذلك أن التبـــاين من المنظور الإحصائي تتعدد مجالاته بصورة ملحوظة، على سبيل المثال فإن التباين يستخدم بكثرة ضمن الانحدار وهناك التباين المتنبأ به والتباين غير المتنبـــأ بـــه (Predicted and unpredicted variance) فتباين الدرجات حول متوسط التوزيع هو تبـــاين يمكـــن تفسيره أو التنبؤ به ، أما تباين الدرجات حول خط الانحدار فهو التباين الذي لا يمكـــن تفسيره أو الذي لا يمكن التنبؤ به

التباين الذي نعنيه في الجزئية هو التباين كأحد مقاييس التشتت عن المتوسط (Mean)، ومن هذا المنظور، فإن التباين يعنى متوسط مربعات الانحراف المعياري، أي أن:

التباين = ع٢

فإذا كان الانحراف المعياري لمتوسط درجات العينة على مقياس مشاهدة التليفزيون هـــو ٥,٥ فإن التباين يكون (٥,٥)^{٢} أي ٣٠,٢٥ وعلى هذا الأســـاس يمكـــن معرفـــة الانحراف المعياري بمعلومية التباين، فإذا كان التباين هو ٣٠,٢٥ فإن الانحراف المعيـــاري هو الجذر التربيعي للتباين

$$\sqrt{٣٠,٢٥} = ٥,٥$$

ويستخدم تحليل التباين في رصد وتحليل معنوية الفروق بين متوسطات درجات عينات أو مجموعات المفحوصين، ومن أبرز التطبيقات في هذا المضمار ما يعرف بتحليل التباين أحادي الاتجاه One Way ANOVA الذي يشيع استخدامه في كافة التخصصات العلمية لمعرفة معنوية الفروق بين متوسطات ثلاث مجموعات فأكثر من حيث (متغير تابع واحد فقط) لنفرض أن لدينا عينة من الأفراد تضم ثلاث مجموعات حسب الحالة الاجتماعية، وهذه المجموعات هي:

- مجموعة المتزوجين
- مجموعة الأعزب
- مجموعة المطلقين

أي أنه حسب (الحالة الاجتماعية) فإن عينه البحث تتضمن المجموعات الثلاث المذكورة، لنفرض أن الدراسة متضمنة تطبيق مقياس استخدام وسائل الإعلام على عينه قوامها ٥٠٠ شخص (N=500) وكانت متوسطات الدرجات والانحرافات المعيارية لمجموعات تلك العينة – حسب متغير الحالة الاجتماعية – على هذا المقياس على النحو المبين بالجدول الآتي:

المجموعة	ن	م	ع
الأعزب	١٠٣	٣٤,٨	٢,٩
المتزوجون	٣٤٧	٣١,٩	٦,٧
المطلقون	٥٠	٢٩,٠٥	٦,٥
العينة	٥٠٠	٣٦,٣٢	٧,٩٢

قبل تطبيق تحليل التباين أحادي الاتجاه يمكن بسهولة التأكد من توافر شروط هذا التطبيق (خاصة اعتدالية التوزيع ، تجانس التباين على النحو الذي سبقت الإشارة إليه عند الحديث عن t test ..)

وكما واضح فإن طبيعة البيانات المبينة بالجدول توضح أن المجموعات مستقلة عن بعضها البعض، وأن لدينا متغيراً تابعاً واحداً (Only one dependent variable) وأن هذا المتغير التابع وحدة قياسه كمية (Quantitative) ويتمثل هذا المتغير في الدرجة على مقياس استخدام وسائل الإعلام . إن التساؤل الذي يجيب عليه تحليل التباين أحادى الاتجاه هو : هل توجد فروق جوهرية بين مجموعات العينة في الدرجة على مقياس استخدام وسائل الإعلام حسب متغير الحالة الاجتماعية ؟ بمعنى أخرى هل يختلف متوسط درجة استخدام وسائل الإعلام من جانب المتزوجين عن الأعزب عن المطلقين؟ وإذا كان هناك اختلاف جوهري، فما هو مكمن هذا الاختلاف الجوهري: هل يوجد بين مجموعة الأعزب ومجموعة المتزوجين؟ أم أنه يوجد بين مجموعة الأعزب ومجموعة المطلقين ؟ أم أنه يوجد بين مجموعة المتزوجين ومجموعة المطلقين؟ إن تحليل التباين أحادي الاتجاه يجيب على هذا التساؤلات.

من خلال تطبيق تحليل التباين أحادي الاتجاه يتم الحصول على معطيات إحصائية متعددة أهمها :-

- عدد مفردات كل مجموعة (N)
- متوسط درجة كل مجموعة (Mean)
- الانحراف المعياري لمتوسط درجة كل مجموعة Standard deviation
- الخطأ المعياري لمتوسط كل مجموعة Standard error

هذا بالإضافة إلى درجة الثقة، والحدين الأعلى والأدنى لدرجة كل مجموعة من مجموعات العينة

وتتيح الحزمة الإحصائية جدولاً يلخص نموذج تحليل التباين، ويكون معنوناً بالمصطلح ANOVA و يتضمن الاستخلاصات الأساسية ممثلة في مصدر التباين Source of variance ، بمعنى التباين بين المجموعات Between groups والتباين داخل المجموعات Within groups والتباين الكلي. فالتباين بين المجموعات يعني التباين القائم بين المجموعات الثلاثة (مجموعة المتزوجين ومجموعة الأعزب ومجموعة المطلقين)، أما التباين داخل المجموعات فهو التباين بين مفردات كل مجموعة، بينما يعني التباين الإجمالي Total

مجمل مجموع التباين داخل المجموعات ومجموع التباين بين المجموعات، كما يتضمن نموذج تحليل التباين مجموع المربعات Sum of Squares ومتوسط المربعات Mean of squares وكذلك درجات الحرية ، وقيمة F و مستوى الدلالة Sig. والذي يعتبر حاسماً في النموذج:

ANOVA

Source of variance	Sum of squares	df.	Mean of squares	F	Sig.
Between groups	1185.111	2	592.55	9.779	. 000
Within groups	30116.089	497	60.596		
Total	3130102	499			

وكما واضح من هذا النموذج فإن درجة الحرية فيما يخص التباين بين المجموعات وكذلك التباين داخل المجموعات قد حددت قيمة متوسط المربعات (إذ إن قيمة متوسط المربعات هي ناتج قسمة مجموع المربعات على درجة الحرية) كما أن قيمة F بلغت ٩,٧٧٩ وهذه القيمة دالة إحصائياً عند مستوى الصفر (Sig. = . 000) وهذا يكشف بوضوح عن أن هناك فروقاً جوهرية بين مجموعات العينة في الدرجة على مقياس استخدام وسائل الإعلام (لاحظ دائماً القيمة الموجودة تحت كلمة .Sig فإذا كانت أقل من ٠,٠٥ أو أقل من ٠,٠١ فهذا يعنى وجود فروق جوهرية)، أما إذا كانت أكبر من ٠,٠٥ فهذا يعنى عدم وجود فروق جوهرية، وبالتالي يتوقف التحليل عند هذا الحد. وبما كان هناك فروقاً جوهرية فهذا يعنى استمرار التحليل لمعرفة أين توجد الفروق الجوهرية :

- هل هذه الفروق توجد بين مجموعة المتزوجين ومجموعة الأعزب ؟
- أم أنها توجد بين مجموعة المتزوجين ومجموعة المطلقين ؟
- أم أنها توجد بين مجموعة الأعزب ومجموعة المطلقين ؟

ولا يعنى ذلك أن الفروق الجوهرية تكون موجودة بين مجموعتين فقط، فقد تكون موجودة بين كل مجموعات العينة ، كأن تكون هذه الفروق بين المجموعة الأولى والمجموعة الثانية، وكذلك بين المجموعة الأولى والمجموعة الثالثة ، وأيضاً بين المجموعة الثانية

والمجموعة الثالثة .... ويتم حسم هذه المسألة من خلال أحد اختبارات المقارنات المتعددة Multi Comparisons test وتتيح الحزمة الإحصائية للعلوم الاجتماعية مجموعة كبيرة من تلك الاختبارات منها :

- LSD
- Bonferroni
- Sidak
- Scheffe
- R-E-G-wf
- R- E – G – wQ
- S – N – K
- Tukey
- tukey's – b
- Duncan
-Hochlerg's Gt2
- Gobriel
- Waller – Duncan
- Dunnet

ويختار الباحث ما يناسب بياناته من تلك الاختبارات، ويلاحظ أن اختبار شيفة Scheffe للمقارنات المتعددة هو الأكثر شيوعاً في بحوث الاتصال والبحوث النفسية والاجتماعية والتربوية. بموجب تنفيذ اختبار المقارنات المتعددة يتم الحصول على معطيات إحصائية متعددة منها مكمن الفروق الجوهرية الموجودة بين متوسطات مجموعات العينة. وفي مثالنا السابق كشف إجراء المقارنات لمتعددة عن المعطيات الموضحة بالجدول الآتي:

المقارنة	مستوى الدلالة
الأعزب –المتزوجون	٠,٠٠٨
الأعزب – المطلقون	صفر
المتزوجون –المطلقون	٠,٠٨٦

ومن هذا الجدول نتبين بوضوح أن مستوى المعنوية بين متوسط درجة مجموعة المتزوجين ومتوسط درجة مجموعة المطلقين يبلغ ٠,٠٨٦ وهذه القيمة أكبر من ٠,٠٥ وبالتالي لا توجد فروق جوهرية بين المتزوجين والمطلقين من حيث الدرجة على مقياس مشاهدة

التليفزيون ($p > 0.05$) وقد أشرنا فيما سبق إلى أن متوسط درجة المتزوجين علـــى هـــذا المقياس هو (٣١,٩) درجة، أما متوسط درجة المطلقين فهو (٢٩,٠٥) درجة . ومن حيث المقارنة بين متوسط درجة مجموعة الأعزب ومتوسط درجة المطلقين على مقيـــاس مشاهدة التليفزيون فإن مستوى الدلالة يساوي الصفر، أي أنه أصغر من ٠,٠٥ وبالتـــالي هناك فروق جوهرية بين مجموعة الأعزب ومجموعة المطلقين، وقد أشرنا فيما سبق إلى أن متوسط درجة مجموعة الأعزب على هذا المقياس هو (٣٤,٨) درجة، أما متوســـط درجة مجموعة المطلقين فهو (٢٩,٠٥) درجة، أي أن الأعزب يـــشاهدون التليفزيـــون بزيادة حقيقية أو جوهرية عن المطلقين($p < 0.05$) ، أما فيما يخص المقارنة بين متوســـط درجة مجموعة الأعزب ومتوسط درجة مجموعة المتزوجين على مقياس مشاهدة التليفزيون فإن مستوى الدلالة يساوي ٠,٠٠٨، أي أنه أصغر كثيراً من ٠,٠٥ وبالتـــالي هنـــاك فروق جوهرية بين مجموعة الأعزب ومجموعة المتزوجين، وقد سبقت الإشـــارة إلى أن متوسط درجة مجموعة الأعزب هو (٣٤,٨) درجة، أما متوسط درجـــة مجموعـــة المتزوجين فهو (٣١,٩) درجة، أي أن مشاهدة التليفزيون تزداد بفروق جوهرية لـــدى مجموعة الأعزب عن مجموعة المتزوجين($p < 0.05$)

التفاعل الثنائي بين المتغيرات:

في الجزئية السابقة كانت المقارنة بين مجموعات العينة مصنفة حسب متغير واحـــد، ألا وهو متغير الحالة الاجتماعية، لكن متوسط الدرجة على مقياس مشاهدة التليفزيون قـــد يختلف حسب متغيرين معاً (وليس حسب المتغيرين كل على حدة)، كأن يختلـــف هـــذا المتوسط حسب متغير الجنس ومتغير الحالة الاجتماعية (معاً) ، وإذا صنفنا عينة الدراســـة حسب هذين المتغيرين يتضاعف عدد مجموعات العينة، فحسب متغير الحالة الاجتماعيـــة بمفرده كان لدينا ثلاث مجموعات (أعزب، متزوج، مطلق)، أما إذا أخذنا بالاعتبار متغير الجنس (ذكور وإناث)، فسوف يكون لدينا ثمان مجموعات هي:

- مجموعة الذكور الأعزب
- مجموعة الإناث الأعزب
- مجموعة الذكور المتزوجين

- مجموعة الإناث المتزوجات
- مجموعة الذكور المطلقين
- مجموعة الإناث المطلقات
- مجموعة الذكور الأرمل
- مجموعة الإناث الأرمل

أي أن كل مجموعة حسب الحالة الاجتماعية تم تقسيمها إلى مجموعتين حسب الجنس، أو أن كل مجموعة حسب الجنس تم تقسيمها إلى أربع مجموعات حسب الحالة الاجتماعية:

الحالة الاجتماعية	**الجنس**	
	ذكور	**إناث**
أعزب		
متزوج		
مطلق		
أرمل		

لنفرض أن التساؤل هو : هل يختلف متوسط مجموعة الذكور الأعزب عن متوسط مجموعة الإناث الأعزب في الدرجة على مقياس مشاهدة التليفزيون؟ يلاحظ من هذا التساؤل أننا أخذنا بالاعتبار متغيرين مستقلين معا، هما متغير الجنس ومتغير الحالة الاجتماعية ، وبالتالي أصبح لدينا ثمان مجموعات هي : الذكور الأعزب، الإناث الأعزب، الذكور المتزوجين، الإناث المتزوجات، الذكور المطلقين، الإناث المطلقات، الذكور الأرامل، الإناث الأرامل.

ويعبر عن ذلك إحصائياً بالتفاعل الثنائي بين المتغيرات المستقلة من حيث تأثير هذا التفاعل على متغير تابع، أي أن التعدد هنا إنما هو على مستوى المتغيرات المستقلة، أما المتغير التابع (The dependent variable) فهو متغير واحد فقط، وهذه النقطة جديرة بالتأكيد عليها لأن التفاعل الثنائي في إطار تحليل التباين أحادي الاتجاه، حيث ينصب التحليل على متغير تابع واحد فقط (وهذا يختلف عن تحليل التباين ثنائي ومتعدد الاتجاهات حيث يكون هناك أكثر من متغير تابع.

وكمثال تطبيقي عن استخدام التفاعل الثنائي بين المتغيرات في إطار تحليل التباين أحادي الاتجاه، نفترض أن إحدى الدراسات عن التوجه نحو المهنة لدى عينة من القائمين بالاتصال (ن= ٩٢٨)، وقد تبين أن استجابات المفحوصين على مقياس التوجه نحو المهنة جاءت بمتوسط قدره ١٠,١ وانحراف معياري ٣,٧ وكانت قيمة متوسط الاستجابات تشكل ٦٧,٣% من مجمل الدرجة، وتختلف قيمة متوسط التوجه نحو المهنة حسب المتغيرات الديموجرافية والمهنية على النحو المبين بالجدول الآتي:

المتغيرات	ن	م	ع	مؤشرات إحصائية
الجنس: ذكور	٢٨٢	٩	٣,٦	T=6.1*
إناث	٦٤٦	١٠,٦	٣,٧	
التعليم: جامعي	٨٣٣	١٠,٢	٣,٨	T=1.4
فوق الجامعي	٩٥	٩,٦	٣,٦	
مدة الخبرة: أقل من ١٠ سنوات	٥٢٤	١٠,١	٣,٦	T=0.06
١٠ سنوات فأكثر	٤٠٤	١٠,٢	٣,٨	
الدخل: منخفض	٣٨٧	١٠,١	٣,٨	F =1.1
متوسط	٢٨٠	١٠,٤	٣,٦	
مرتفع	٢٦١	٩,٩	٣,٨	
العينة	٩٢٨	١٠,١	٣,٧	

*P<0.05

واضح من هذا الجدول وجود فروق بين مجموعات العينة من حيث التوجه نحو المهنة، وإن كانت هذه الفروق غير ذات دلالة إحصائية فيما يخص متغير مدة الخبرة ومتغير المستوى التعليمي (P>0.05) ، لكن هناك فروقاً جوهرية في التوجه نحو المهنة لدى مجموعات

العينة حسب متغير الجنس (P<0.05) ، حيث يتضح من الجدول أن الإناث أكثر توجهاً نحو المهنة مقارنة بالذكور (م=١٠,٦ للإناث مقابل ٩ للذكور)، وإذا كان ذلك فيما يخص معنوية الفروق بين مجموعات العينة في التوجه نحو المهنة حسب المتغيرات الموضحة بالجدول كل متغير على حدة، فإنه باستخدام التفاعل الثنائي(Bilateral Interaction) بين هذه المتغيرات خلصت الدراسة إلى النتيجة الموضحة بالجدول الآتي:

مصدر التباين	درجة الحرية	مجموع المربعات	متوسط المربعات	حجم التأثير	قيمة F
الجنس× التعليم	١	١١,٤	١١,٤	٠,٠٠١	٠,٨٥
الجنس×مدة الخبرة	١	٣,٦	٣,٦	صفر	٠,٣
الجنس× الدخل	٢	٦١,٥	٣٠,٧	٠,٠٠٥	٢,٣
التعليم ×مدة الخبرة	١	٨٨,٧	٨٨,٧	٠,٠٠٧	٦,٤*
التعليم × الدخل	٢	١٦	٨	٠,٠٠١	٠,٦
مدة الخبرة×ا لدخل	٢	٦٢,٦	٣١,٣	٠,٠٠٥	٢,٤

*P<0.05

يتضمن هذا الجدول ستة تفاعلات ثنائية هي كافة التفاعلات الثنائية الممكنة بين المتغيرات (وذلك على أساس أن لدينا أربعة متغيرات مستقلة ، وبالتالي يكون عدد التفاعلات الثنائية الممكنة هو ٤×٣ ÷٢= ٦ ومن بين هذه التفاعلات الستة يوجد تفاعل واحد فقط ذو دلالة إحصائية للدرجة على مقياس التوجه نحو المهنة، ألا وهو التفاعل بين متغير المستوى التعليمي ومتغير مدة الخبرة، ونظراً لوجود دلالة لهذا التفاعل، فإنه يتعين تقصي متوسطات درجات مجموعات العينة حسب هذين المتغيرين معاً (متغير المستوى التعليمي ومتغير مدة الخبرة)، وباستخدام التفاعل ثنائي الاتجاه جاءت النتيجة على النحو المبين بالجدول الآتي:

التعليم	مدة الخبرة			
	أقل من ١٠ سنوات		١٠ سنوات فأكثر	
	م	ع	م	ع
جامـــــعي	١٠,٢	٣,٧	١٠,١	٣,٨
فوق الجامعي	٨,٦	٣,٤	١٠,٥	٣,٦

يتضمن الجدول أربع مجموعات فرعية تضمها العينة حسب متغير مدة الخبرة ومتغير التعليم معا:

- مجموعة المفحوصين ذوي التعليم الجامعي والخبرة الأقل من ١٠ سنوات
- مجموعة المفحوصين ذوي التعليم فوق الجامعي والخبرة الأقل من ١٠ سنوات
- مجموعة المفحوصين ذوي التعليم الجامعي والخبرة ١٠ سنوات فأكثر
- مجموعة المفحوصين ذوي التعليم فوق الجامعي والخبرة ١٠ سنوات فأكثر

وقرين كل مجموعة متوسط الدرجة (م) والانحراف المعياري (ع)، ومن الجـــدول نتـــبين بوضوح أن القائمين بالاتصال ذوي الخبرة الأطول في العمل، الحاصلين علـــى مـــؤهلات فوق الجامعية هم أكثر مجموعات العينة تعبيراً عن التوجه الإيجابي نحو المهنة (م=١٠,٥)، يليهم مجموعة القائمين بالاتصال ذوي الخبرة الأقل من عشر سنوات الحاصلين على مؤهل جامعي (م= ١٠,٢) وفي المرتبة نفسها تقريباً تأتي مجموعة القائمين بالاتصال ذوي الخبرة الأكثر من عشر سنوات الحاصلين على مؤهل جامعي (م= ١٠,١) ، أما أقل المجموعات تعبيراً عن هذا التوجه – حسب مدة الخبرة والمستوى التعليمي – فهي مجموعة القـــائمين بالاتصال ذوي الخبر الأقل من عشر سنوات الحاصلين على مـــؤهلات فـــوق الجامعيـــة (م=٨,٦).

هكذا يتضح لنا أن استخدام التفاعل ثنائي الاتجاه قد مكننا من رصد متوسطات درجات مجموعات العينة على مقياس التوجه نحو المهنة وذلك حسب مدة الخبرة والمستوى التعليمي معاً.

ويمكن استخدام التفاعل متعدد الاتجاه (Multi interaction) لرصد متوسطات درجات مجموعات العينة حسب أكثر من متغيرين مستقلين أو تصنيفيين (كأن يتم التعرف على متوسطات درجات مجموعات القائمين بالاتصال على مقياس التوجه نحو المهنة حسب متغيرات الجنس والمستوى التعليمي ومدة الخبرة معاً)، وحسب هذه المتغيرات الثلاثة يكون لدينا المجموعات الآتية:

الجنس	ذكور		إناث	
التعليم:	جامعي	فوق الجامعي	جامعي	فوق الجامعي
الخبرة: أقل من ١٠				
١٠ فأكثر				

- مجموعة الذكور ذوي التعليم الجامعي ، ذوي الخبرة الأقل من ١٠ سنوات
- مجموعة الذكور ذوي التعليم الجامعي ، ذوي الخبرة ١٠ سنوات فأكثر
- مجموعة الذكور ذوي التعليم فوق الجامعي ، ذوي الخبرة الأقل من ١٠ سنوات
- مجموعة الذكور ذوي التعليم فوق الجامعي ، ذوي الخبرة ١٠سنوات فأكثر
- مجموعة الإناث ذوات التعليم الجامعي ، ذوات الخبرة الأقل من ١٠ سنوات
- مجموعة الإناث ذوات التعليم الجامعي ، ذوات الخبرة ١٠ سنوات فأكثر
- مجموعة الإناث ذوات التعليم فوق الجامعي ، ذوات الخبرة الأقل من ١٠ سنوات
- مجموعة الإناث ذوات التعليم فوق الجامعي ، ذوات الخبرة ١٠ سنوات فأكثر

أي أنه حسب متغيرات الجنس والتعليم ومدة الخبرة يكون لدينا ثمان مجموعات فرعية، ومن الطبيعي أن يزداد عدد مجموعات العينة كلما زاد عدد المتغيرات التي ستؤخذ بالاعتبار في التفاعل ونحن بصدد تطبيق تحليل التباين أحادي الاتجاه، كما أن زيادة عدد المتغيرات التي يتضمنها التفاعل يتطلب أن تكون العينة كبيرة.بما يضمن وجود العدد الكافي من المفردات في كل مجموعة وفي الوقت نفسه يضمن عدم الإخلال بشرط أساسي في استخدام المقارنات ألا وهو التقارب بين عدد المفردات المجموعات

التي تتم المقارنة بينها، فلا يعقل مثلاً أن تكون إحـــدى المجموعـــات ١٠٠ مفـــردة ومجموعة أخرى خمس مفردات

الفصل الخامس

مدخل تعريفي بالتحليل العاملي

تقديم:

يتضمن هذا الفصل تعريفاً موجزاً بالتحليل العاملي من حيث مفهومه وخصائصه ومفاهيمه الأساسية واستخداماته. كما يتطرق الفصل إلى طرق تدوير العوامل وتقدير درجاها وتسميتها وتفسيرها، وكذلك أساليب التحليل العاملي المعاصرة. ويتضمن الفصل جانباً تطبيقياً يتعلق بتصميم وإجراء الدراسة العاملية وصياغة نتائجها مع الإشارة إلى استخدام الحاسوب في التحليل العاملي، وينتهي الفصل بالإشارة إلى أهم الأخطاء الشائعة في التحليل العاملي باعتباره من أهم الأساليب الإحصائية في البحث العلمي.

أولاً: مفهوم التحليل العاملي وخصائصه:

يعتبر التحليل العاملي (Factor Analysis) من الأساليب الإحصائية المهمة التي يمكن أن يستخدمها الباحث في تحليل البيانات متعددة المتغيرات، ودراسة العلاقات القائمة بين تلك المتغيرات واختزالها في عدد أقل من العوامل(Factors) التي يمكن أن تفسر الظاهرة مجال الدراسة، فمن خلال التحليل العاملي يتمكن الباحثون من دراسة الظواهر المتشابكة التي تتسم بالتعقيد وتعدد المتغيرات في محاولة صياغة نتائج هذا التحليل في أطر نظرية علمية، ومن هنا يعتبر التحليل العاملي طريقة منهجية وأسلوب تحليل إحصائي وبنية نظرية، حيث يسمح بالتعامل مع البيانات الكمية والنوعية بكل من الطريقتين الاستقرائية والاستنباطية، كما يلعب التحليل العاملي دوراً مهماً في البحث العلمي، فجميع العلوم تهدف إلي بناء وتطوير نظريات تفسر العلاقات القائمة، علي الرغم من أنه في نطاق أي مجال علمي ربما تقتصر بعض الدراسات علي وصف الظواهر عن طريق تلخيص البيانات بحيث يمكن فهم العلاقات الإمبيريقية بين متغيراتها.

والهدف من استخدام التحليل العاملي يتعلق عادة بتلخيص العلاقات بين المتغيرات بطريقة دقيقة ومنظمة ومقتصدة من أجل فهم أفضل وخلق تصور فكري أوضح للظواهر التي يهتم الباحثون بدراستها، فعن طريق التحليل العاملي يمكن اختزال عدد كبير من المتغيرات واستخلاص أكبر قدر من المعلومات منها لأن المجموعة الأصلية من المتغيرات المتعددة يتم تجميعها في عدد قليل من العوامل التي تفسر التباين في هذه المجموعة الكبيرة من المتغيرات. وعلي الرغم من أن التحليل العاملي له بعض الخصائص التي يشترك فيها مع أساليب إحصائية أخرى، ، إلا أنه يتميز عن هذه الأساليب الأخرى بخصائص متعددة نذكر منها ما يلي:

(١) يتناول التحليل العاملي مجموعات كبيرة من البيانات المستمدة من الاختبـارات والمقاييس بمختلف أنواعها، مثل الاستبيانات ودراسات الحالات والمسوح، حيث يمكن إجراء التحليل العاملي علي عشرات المتغيرات خاصة من خلال البـرامج الإحصائية الحاسوبية المتطورة

(٢) يعد التحليل العاملي من الأساليب الإحصائية التي تتميز بالمرونة حيـث يمكـن توظيفه في تصميمات بحثية متعددة، وذلك للتحقق من صحة الفروض، ورسـم خرائط المفاهيم، ودراسات الحالات، والدراسات الطولية، كما يمكن الإفادة منه في تحليل بيانات متنوعة، مثل المسوح الاجتماعية، والسلاسل الزمنية، ودرجـات الاختبارات العقلية ومقاييس الشخصية والجوانب النفسية الحركيـة، والبيانـات المتعلقة بالانتخابات والرأي العام، وغير ذلك من المجالات الحيوية.

(٣) يمكن باستخدام التحليل العاملي دراسة الظواهر المختلفة ميدانياً دون الحاجـة إلي إجراءات مختبريه صارمة لضبط المتغيرات الدخيلة، وبذلك يتم تحليـل العلاقـات القائمة بين المتغيرات وواقع البيئة الفعلية إلي وحدات أو أنماط مستقلة من السلوك وتحديد الآثار المستقبلية.

(٤) يعد التحليل العاملي من الأساليب الإحصائية متعددة المداخل، فعلي الرغم من أن جذوره ممتدة في العلوم الاجتماعية وبخاصة علم النفس، إلا أن علماء الرياضيات ومناهج البحث وغيرهم قد تناولوه بفيض من الدراسات المتخصصة، وقد نـال قدراً من الاهتمام في مراجع مناهج البحث أكثر من الأساليب الإحصائية الأخرى المستخدمة في العلوم الاجتماعية.

(٥) أسهمت أساليب التحليل العاملي في إيجاد تكامل بينها وبين كثير من الأسـاليب الإحصائية الأخرى متعددة المتغيرات المتعلقة بالارتباطات، مثل الانحدار المتعـدد، وتحليل المسارات، والارتباط الجزئي وشبه الجزئي، والارتباط المركب والمـوازين متعددة الأبعاد وتحليل التباين، وغيرها.

(٦) يؤدي التحليل العاملي إلي مجموعة من المعادلات التي يمكن استخدامها في وصف الظواهر والتنبؤ بها، حيث تستخدم هذه المعادلات في تطوير النظريات والتوصـل إلي استنباطات لم تكن واضحة.

(٧) يمكن باستخدام التحليل العاملي تصنيف الأفراد والأشياء في أنماط متمايزة عن طريق تحليل العلاقات وذلك للتوصل إلي تجمعات من الأفراد المتماثلين أو الأشياء المتشابهة من مجموعات مختلفة.

(٨) يتميز التحليل العاملي بأنه يسمح بالتصور البصري للعلاقات بين المتغيرات المتعلقة بالظواهر المختلفة عن طريق التمثيل الهندسي، كما يسمح ببناء نماذج مجردة للواقع الاجتماعي والظواهر المختلفة من خلال المعادلات والتمثيلات الهندسية، لذلك يطلق عليه البعض "حساب العلوم الاجتماعية Calculus of Social Sciences".

ثانياً:المفاهيم الأساسية في التحليل العاملي:

هناك مجموعة من المفاهيم الأساسية التي لا بد أن يعرفها الباحث حتى يحقق فهماً أفضل للتحليل العاملي،علما بأن الاستيعاب الإحصائي لتلك المفاهيم يتطلب المعرفة بأساسيات الرياضيات العالية(خاصة الجبر الخطي وهندسة المتجهات والتفاضل)، لكننا هنا سوف نقدم تعريفاً موجزاً بأهم مفاهيم التحليل العاملي بما يساعد الباحثين على ما تعنيه معطيات التحليل العاملي:

(١) العامل:

إن مفهوم العامل (Factor) مفهوم جوهري في التحليل العاملي، ولكي نوضح المقصود بهذا المفهوم نقدم المثال التالي:

إذا حللنا العدد ٦ إلى عوامله الأولية، فإننا نحصل على المعادلة التالية:

٦= ٣×٢×١

إن الأعداد ٣، ٢، ١ هي عوامل العدد ٦ أو المكونات الرئيسية للعدد ٦ فإذا كان الرقم ٦ يدل مثلاً على المساحة، فإن العدد ٣ يدل على الطول، أما العدد ٢ فيدل على العرض، ولا يدل العدد ١ على شيء. لكن عندما يكون الرقم ٦ يدل على الحجم فإن العدد ٣ يدل على الطول، أما العدد ٢ فيدل على العرض، وقد يدل العدد ١ على الارتفاع. في هذا المثال نتبين أن الرقم ٦ يعني شيئاً محدداً، وأن الأعداد ٣، ٢، ١ هي مكونات هذا الشيء .

ومن المنظور البحثي فإن العامل (Factor) هو المعنى العام الذي يجمع أكثر من بند أو متغير، ومن المنظور الإحصائي فإن العامل يلخص الارتباطات (Correlations) بين البنود

أو المتغيرات، أي أن العامل Factor صيغة رياضيا كبديل عن هــذه البنــود أو تلــك المتغيرات، وقد كشفت دراسات سبيرمان هذه الحقيقة عندما أوضحت أن "العامل" هو السبب المباشر وراء وجود الارتباط الموجب القائم بين أي ظاهرتين، فــإذا فرضــنا أن الظاهرة (أ) ترتبط ارتباطاً موجباً بالظاهرة (ب)، فإن هذا الارتباط سببه وجود عامــل مشترك (ش) يؤثر تأثيراً موجباً في الظاهرتين (أ) و (ب)، وعندما يختفي تأثير العامل (ش) يتلاشى الارتباط بين هاتين الظاهرتين. وقد تطور مفهوم العامل فيما بعد، ليعني الارتباط بين أي عدد من الظواهر أو المتغيرات

إن العوامل Factors التي يهدف التحليل العاملي الكشفي لاستخلاصها مــن مــصفوفة الارتباطات بين المتغيرات تعد بمثابة تكوينات فرضية Hypothetical Constructs تنطوي علي المتغيرات، وهناك أنواع مختلفة من العوامل لكل منها تباين خاص به، وتتمثل تلــك العوامل في العامل العام، العوامل المشتركة، العوامل النوعية.

فالعامل العام General Factor هو العامل الذي تتشبع به جميع المتغيرات، فإذا كان لدينا استبانة تتكون من عشرين بنداً لقياس الحاجات التي يشبعها الترفيه لــدى الجمهور،فــإن العامل العام هو ذلك العامل الذي تتشبع به جميع بنود الاستبانة.

أما العوامل المشتركة،فكل عامل منها تتشبع به مجموعة من البنود،كأن نتــبين أن البنــود أرقام:١، ٢، ٣، ٤، ٥، ٦،٧ ، ١٢، ١٨ تتشبع بالعامل المشترك الأول، أما البنود أرقــام ٨، ٩، ١٠، ١١، ١٣،١٩ ، ٢٠ فتتشبع بالعامل المشترك الثاني، أما العامل المشترك الثالث فتتشبع به البنود١٧،١٥،١٦،١٤ .

وفي التحليل العاملي تكون تشبعات كل متغير بالعوامل موجبة أو سالبة، وربمــا تكــون بعض هذه التشبعات منخفضة، ونقصد بالتشبع العاملي درجة ارتباط المتغير بالعامل. وهنا ينبغي أن نقرر بطريقة اعتبارية حدود التشبعات المنخفضة التي تعد غير ذات أهمية، وعدد هذه التشبعات بالنسبة لعامل معين لكي نعتبره عاملاً مشتركاً وليس عاملاً عاماً.

أما فيما يخص العوامل النوعية Specific Factors فهي العوامل التي يتشبع بها بند واحد فقط،فإذا كان هناك أحد العوامل الذي يتشبع به بند واحد فإن هذا العامل يطلق عليه عامل نوعي

وفي التحليل العاملي نجد أن العوامل المستخلصة تفسر نسبة معينة من التباين، مثال ذلــك أن يسفر التحليل العاملي لبيانات مستمدة من تطبيق مقياس استخدامات وسائل الإعــلام

عن خمسة عوامل ، وتبلغ قيمة التباين الذي تفسره تلك العوامل ٠,٥٨ وهذا يعني أن العوامل المستخرجة تفسر ٥٨% من التباين وأن ٤٢% تفسرها متغيرات أخرى، أي أن العوامل المستخلصة لا تفسر التباين الكلي كاملاً Total Variance .وينبغي أن يكون مربع تشبعات بند معين بالعوامل مساوياً لتباين ذلك المتغير، فإذا كان التباين يساوي الواحد الصحيح في مصفوفة معاملات الارتباط، فإن جميع مربعات تشبعات العامل لهذا المتغير يجب أن تكون مساوية الواحد الصحيح.

(٢) مصفوفة الارتباط:

في التحليل العاملي تكون مصفوفة الارتباط (Correlation Matrix) معبرة عن قيم معاملات الارتباط بين المتغيرات بعضها بعضاً أو العوامل بعضها بعضاً أو بين العوامل والمتغيرات.

ونقطة البدء في التحليل العاملي تكوين مصفوفة الارتباطات بين مجموعة المتغيرات التي عددها (ن)، فإذا كان الاستبيان يقيس أربعة أبعاد هي:

- علاقة الجمهور بالتليفزيون
- علاقة الجمهور بالراديو
- علاقة الجمهور بالصحف اليومية
- علاقة الجمهور بالصحف الأسبوعية

فإن مصفوفة معاملات الارتباط تكون على النحو المبين بالجدول الآتي:

الأبعاد	مشاهدة التليفزيون	سماع الراديو	قراءة الصحف اليومية	قراءة الصحف الأسبوعية
مشاهدة التليفزيون	١	٠,٣٩	٠,٥٨	٠,٣٨
سماع الراديو	٠,٣٩	١	٠,٣٥	٠,٣٦
قراءة الصحف اليومية	٠,٥٨	٠,٣٥	١	٠٥٣
قراءة الصحف الأسبوعية	٠,٣٨	٠,٣٦	٠٥٣.	١

أي أن المتغيرات توضع علي كل من البعدين الأفقي والرأسي، وقيم معامل الارتباط بين كل متغيرين منها في خلايا المصفوفة. كما يلاحظ أن الجدول منقسم إلى مثلثين مشتركين في الوتر، كل مثلث يتضمن الخلايا نفسها،ويفصل بين المثلثين الخلايا التي بها الرقم (١)، أي أن مصفوفة الارتباط في المثلث العلوي تشتمل علي جميع الارتباطات بين المـــتغيرات، كما أن هذه الارتباطات نفسها تتضمنها مصفوفة الارتباط في المثلث السفلي لتيسير جمع القيم في أي صف وأي عمود في عمليات استخلاص العوامل.

(٣) الاشتراكيات أو قيم الشيوع:

يُقصد بالاشتراكيات في مصفوفة الارتباطات مجموع مربعات ارتباطات المتغير بنفسه الذي يرجع إلي عامل مشترك أو عوامل مشتركة، وبذلك تدل علي نسبة التباين في المتغير والتي تُعزي إلي -أو تُفسر بواسطة- هذا العامل أو العوامل المشتركة. وينبغي هنا التمييز بـــين الاشتراكيات Communalities وبين معامل ثبات Reliability درجات متغير معـــين. فالثبات يدل علي ارتباط المتغير بنفسه الذي يُعزي إلي العوامل النوعية المتعلقة بالاختبـــار الذي نريد التحقق من ثباته،وكذلك العوامل التي يشترك فيها مع غيره من العوامل. أمـــا الخلايا القطرية في المصفوفة التي تُدون فيها الاشتراكيات فتترك خالية في بـــدء التحليـــل العاملي، وينبغي تقدير هذه القيم أو تخمينها عند بدء العمليات الحسابية، لذلك فإن قـــيم الاشتراكيات لا تكون دقيقة تماما، حيث إنه دائما يتم تقديرها.

وبعد وضع قيم الاشتراكيات في الخلايا القطرية تبدأ عمليات استخلاص العوامـــل مـــن مصفوفة الارتباطات بين المتغيرات، وذلك بجمع قيم الارتباطات المدون في كل عمـــود، وقسمة مجموع كل عمود علي الجذر التربيعي للمجموع الكلي للأعمدة، وذلك لتيـــسير المقارنة بين مجاميع الأعمدة، وبينما تدل الاشتراكيات علي التباين المـــشترك Common Variance، فإن التباين النوعي الخاص بالمتغير Unique Variance يـــساوي الواحـــد الصحيح مطروحا من الاشتراكيات.

(٤) التشبعات

يُقصد بالتشبع (Saturation) للمتغير ارتباط ذلـــك المـــتغير بعامـــل Factor معـــين تم استخلاصه، وتُفسر قيمة التشبع بنفس طريقة تفسير معامل ارتباط بيرسون، حيث يـــتم تربيع قيمة التشبع العاملي لكل من المتغيرات في مصفوفة التشبعات للحصول علي نـــسبة التباين في المتغير التي يمكن تفسيرها بواسطة العامل المـــستخلص. ومتوســـط مربعـــات التشبعات العاملية في هذه المصفوفة يدل علي مقدار التباين الكلي في المتغيرات كمجموعة

والتي يتم تفسيرها بواسطة هذا العامل أو ذاك، أما متوسط مجموع مربعات التـــشبعات العاملية في جميع الأعمدة فيدل علي نسبة التباين الذي يمكن تفسيره بواســطة العوامـــل المستخلصة، وهذا يكون مؤشراً لمدى صلاحية هذه العوامل في تفـــسير تبـــاين مجموعـــة المتغيرات الأصلية.

وينبغي ملاحظة أن مجموع مربعات التشبعات في أعمدة مـــصفوفة التـــشبعات العامليـــة يساوي مجموع اشتراكيات المتغيرات، ومجموع نسب التباين في المـــتغيرات الـــتي يمكـــن تفسيرها بواسطة العوامل المستخلصة يساوي متوسط قيم الاشتراكيات. وكلما كان المتغير مشتركا بدرجة أكبر مع غيره من المتغيرات في العوامـــل المـــشتركة؛ زاد مربـــع قيمـــة الاشتراكيات. ونظرا لأن عدد العوامل المستخلصة يكون عادة أقل من عـــدد المـــتغيرات الأصلية، فإننا لا نستطيع الحصول علي قيم المتغيرات باستخدام العوامل، وإنما يمكن تقدير قيمها باستخدام معادلة الانحدار المتعدد لكل متغير علي هذه العوامل. فـــإذا افترضـــنا أن العوامل غير متعامدة، أي غير مرتبطة، فإن مجموع مربعات التشبعات في أي صف مـــن صفوف مصفوفة التشبعات العاملية يكون مساويا مربع معامل الارتباط المتعدد بين العوامل المستخلصة ومتغير معين من مجموعة المتغيرات. ولتوضيح ذلك نفترض أن لدينا مـــصفوفة التشبعات العاملية التالية التي تم استخلاصها باستخدام أحد أساليب التحليـــل العـــاملي، وذلك لثمانية متغيرات، وكذلك اشتراكياقما كما في الجدول التالي:

المتغيرات	العوامل قبل التدوير		
	العامل الأول	العامل الثاني	الاشتراكيات
١	٠,٦٥	٠,٥٧	٠,٧٥
٢	٠,٥٤	٠,٥٤	٠,٥٨
٣	٠,٦١	-٠٤٥.	٠,٥٧
٤	٠,٦٣	-٠,٥٤	٠,٦٩
٥	٠,٥٦	٠,٥٤	٠,٦١
٦	٠,٧٢	-٠,٥٩	٠,٨٧
٧	٠,٦٧	-٠,٤٥	٠,٦٥
٨	٠,٦٤	٠,٦٠	٠,٧٧
التباين المشترك	**٣,١٧**	**٢,٣١**	**٥,٤٩**
نسبة التباين الكلي	**٠,٤٠**	**٠,٢٩**	**٠,٦٩**
نسبة التباين المشترك	**٠,٥٨**	**٠,٤٢**	**١,٠٠**

فالمتغير الأول مثلاً يرتبط بالعامل الأول بمقدار ٠,٦٥ كما يرتبط بالمتغير الثاني بمقدار ٠,٥٧ أي أن تشبع المتغير الأول بالعامل الأول يساوي ٠,٦٥ أما تشبع المتغير الأول بالعامل الثاني فهو ٠,٥٧ وتدل الأرقام المدونة في خانة الاشتراكيات على نسبة التباين الكلي في المتغير، هذا التباين الذي يمكن تفسيره بالعاملين ، على سبيل المثال، فإن العامل الأول والعامل الثاني يفسران ٧٥% من التباين في المتغير الأول، أما النسبة المتبقية وقدرها ٢٥% فإنها تفسر بعوامل أخرى وكذلك بأخطاء القياس، ويلاحظ أن التباين المشترك للعامل الأول يساوي ٣,١٧ (وهذه القيمة هي مجموع مربعات تشبعات المتغيرات بالعامل الأول)، أما عن متوسط هذه القيمة فيكون بقسمتها على عدد المتغيرات، أي ٣,١٧ ÷ ٨ = ٠,٤٠ وهي نفسها قيمة نسبة التباين الكلي التي تمثل نسبة التباين التي يفسرها العامل الأول في المتغيرات الثمانية

أما العامل الثاني فيفسر نسبة ٠,٢٩ من التباين في المتغيرات الثمانية، وبالتالي فإن العاملين الأول والثاني يفسران ٠,٦٩ من التباين الكلي. يلاحظ من الجدول أيضا أن كل قيمة من قيم الاشتراكيات الخاصة بكل متغير تساوي مجموع مربع تشبع المتغير بالعامل الأول + مجموع مربع تشبع المتغير بالعامل الثاني، خذ مثلاً المتغير الأول تجد أن قيمة الاشتراكيات =٠,٧٥ وهذه القيمة تساوي $(٠,٦٥)^٢ + (٠,٥٧)^٢$

ثالثاً: استخدامات التحليل العاملي:

هناك استخدامان رئيسيان للتحليل العاملي، الأول استكشافي، والثاني تأكيدي،وسوف نوضح هذين الاستخدامين فيما يلي:

(أ) **التحليل العاملي الاستكشافي** Exploratory Factor Analysis:

تعد أساليب التحليل الكشفي للبيانات Exploratory Data Analysis من الأساليب الإحصائية التي تهتم بفحص مجموعات البيانات من زوايا متعددة، وهذا التحليل الكشفي وما يزودنا به من معلومات- يمكن أن يؤدي إلي تحليلات تالية أكثر دقة مما يسهم في إلقاء مزيد من الضوء علي الظاهرة التي نهتم بدراستها. وقد يصل الباحث إلى نتائج غير متسقة، الأمر الذي يستفيد منه الباحث في إجراء مزيد من التحليل، وقد يكرر تجربته أو يقسم عينة دراسته إلي مجموعتين أو أكثر، ثم يجمع بياناته، ويقوم بتحليل بيانات كل مجموعة منها علي حدة. فالنتائج المتسقة عبر المجموعات هي التي ينبغي أن يوليها الباحث مزيدا من الدراسة؛ نظراً لأنه يستطيع عن طريقها التحقق من نموذج نظري معين، ومعرفة ما إذا كانت هناك متغيرات كشفية لم تؤخذ بعين الاعتبار أو أقحمت بعض المتغيرات التي لا تتعلق بالدراسة،وبذلك يكون التحليل الكشفي للبيانات عملية تفاعل بين الباحث والأسلوب الإحصائي المستخدم والبيانات، وبالطبع لا يحدث هذا التفاعل إلا إذا كان لدى الباحث معرفة جيدة بالأسلوب الإحصائي الذي يستخدمه وطبيعة البيانات التي يقوم بتحليلها، وكذلك فهم مستبصر للظاهرة التي يهتم بدراستها.

غير أن الباحث ربما لا يكون لديه فكرة واضحة عن عدد الأبعاد التي تنطوي عليها مجموعة معينة من المتغيرات، وهنا يكون التحليل العاملي أسلوباً مناسباً للكشف عن أقل

عدد من العوامل الافتراضية التي تفسر التغاير في البيانات، وبذلك يسهم في الكشف عن إمكانيات اختزال هذه البيانات، ويلاحظ أن الكثير من استخدامات وتطبيقات التحليل العاملي تركز علي هذا الجانب الكشفي للتحليل العاملي Exploratory Factor Analysis، وبخاصة إذا كان المجال البحثي يشتمل علي علاقات معقدة لم تحاول الدراسات السابقة تناولها بطريقة منظمة، وكذلك إذا لم يستطع الباحث التعامل مع المتغيرات باستخدام المنهج التجريبي، فإنه يمكنه أن يتعامل مع الظواهر المعقدة في مواقف واقعية، وبذلك يصبح التحليل العاملي بديلاً عن التجارب المختبرية لأنه يمكن الباحث من تحليل العلاقات بين المتغيرات، والفصل بين المصادر المختلفة للتباين، وضبط المتغيرات التي يهتم بدراستها.

(ب)**التحليل العاملي التأكيدي** Confirmatory Factor Analysis:

علي الرغم من أهمية التحليل العاملي الاستكشافي وشيوع استخدامه، إلا أن هذا لا يعني بالضرورة اقتصار هذا التحليل علي كشف الأبعاد التي تنطوي عليها مجموعة من المتغيرات المرتبطة، فنتائج هذا التحليل ربما تكشف عن تكوينات فرضية Constructs أو فرضيات Hypotheses تتطلب المزيد من البحث. ويعتمد إسهام التحليل العاملي الكشفي في تطوير النظريات وبناء النماذج علي متابعة نتائجه الكشفية بالبحث والدراسة من أجل تأكيد أو رفض فرضية معينة. أي أن التحليل العاملي يستخدم في هذه الحالة للتحقق من صحة الفرضيات Hypothesis Testing. فمثلاً ربما يفترض الباحث أن هناك عاملين أو بعدين مختلفين تنطوي تحتهما مجموعة من المتغيرات، وأن بعض المتغيرات ينتمي إلى العامل الأول، والبعض الآخر ينتمي إلى العامل الثاني، فإذا استخدم التحليل العاملي للتحقق من صحة هذا الافتراض فإنه يطلق عليه في هذه الحالة "التحليل العاملي التأكيدي Confirmatory Factor Analysis". غير أن التمييز بين الاستخدامين الكشفي والتأكيدي لا يكون دائما تمييزاً قاطعاً، لأن الكثير من البحوث تجمع بين الاستخدامين؛ وذلك لأنه يندر أن يقوم الباحث بإجراء التحليل العاملي علي مجموعة عشوائية من المتغيرات، بل يكون مدركا بدرجات متفاوتة طبيعة المتغيرات، والأبعاد التي تضم هذه المتغيرات، فمن الممكن مثلاً أن يفترض الباحث وجود عاملين، ولكنه لا يعرف علي وجه

الدقة أي المتغيرات تمثل كلاً منهما، أو ربما يستخدم الباحث نصف عدد أفـــراد العينـــة للكشف عن البنية العاملية المحتملة، ثم يستخدم النصف الآخر في التحقق من هذه البنيـــة، كما يمكن استخدام التحليل العاملي كأسلوب تنقيبي Heuristic Device، فمثلاً إذا بينت الدراسات السابقة أن الذكاء الإنساني يشتمل علي سبعة عوامل أولية مختلفـــة، ولكـــن الباحث يود التأكد من أن متغيراً معيناً يندرج تحت أحد هذه العوامل، فعندئـــذ يمكـــن استخدام التحليل العاملي في التحقق مما يعنيه هذا المتغير. وكذلك إذا افترض الباحـــث أن نطاق مجموعة من البيانات له جانبان متمايزان أحدهما كمي والآخر نوعي، فإنه يمكنـــه التحقق من صحة هذا الفرض باستخدام التحليل العاملي التأكيدي، فإذا أكدت معطيات التحليل العاملي هذا الفرض، فإن العوامل المختلفة سوف تعكس التمايز النوعي، أمـــا إذا افترض الباحث أن أحد المتغيرات مرتبط بعامل معين بدرجة أكبر من ارتباطه بعامل آخر، فإن هذا التمايز الكمي يمكن التحقق منه أيضاً، وقد يكشف التحليل العاملي عن وجـــود هذا التمايز أو عدم وجوده.

رابعاً: أنماط التحليل العاملي:

من المعروف أن التحليل العاملي لا يقتصر علي تحليل الارتباطات بين المتغيرات التي جمعت بياناتها من عينات من الأفراد، حيث إن تطبيقاته وأدواته أكثر اتساعاً وثراءً، وربما يرجـــع ذلك إلى تعدد أبعاد تغاير البيانات وما تمثلها من ظواهر،وبالتالي فإن أي ظـــاهرة يمكـــن وصفها في ثلاثة أبعاد هي:

- المتغيرات Variables: وهذه تتعلق بالخصائص أو السمات أو السلوك الذي يصف الأفراد أو الأشياء أو القضايا ويميز بين هؤلاء الأفراد أو هذه الأشـــياء،وتتحدد هـــذه المتغيرات بعمليات الملاحظة أو القياس.

- الأفراد Individuals: وهذه تشمل الأفراد والجماعات والأشـــياء الـــتي نقـــيس خصائصها، لذلك يُطلق عليها هورست Horst "الكينونات Entities". كما يمكن أن تشمل القضايا التي نصفها، مثل التضخم السكاني وما يتعلق به من خصائص كالتوزيع الجغرافي للسكان، ومخاطر التضخم، وإسقاطاته، وهكذا.

- المواقف أو الظروف Occasions: وهذه تتعلق بزمن حدوث الظاهرة ومكانها، أي الظروف المتعلقة بالظاهرة.

وتختلف أنماط التحليل العاملي باختلاف وجهي أو بعدي البيانات اللذين نتناولهما، وعلى الرغم من تعدد أنماط التحليل العاملي، إلا أن هناك ثلاثة أنماط هي الأكثـــر اســـتخداما، وتتمثل في:

(١) التحليل العاملي للسمات في مجتمعات من الأفراد، ويطلق علي هذا الـــنمط R-Factor Analysis: وهذا النمط شائع الاستخدام في مختلف مجالات البحـــوث، وفيه يكون لدينا مصفوفة بيانات تشتمل علي (م) من الأفراد أو الكينونات، (ن) من المتغيرات، أي (م × ن)، فعندئذ تكون مصفوفة الارتباطات بين المتغيرات مثني (ن × ن). ومثال ذلك التحليل العاملي لبطارية من الاختبارات أو المقاييس الـــتي تطبق علي عينة من الأفراد في وقت معين، أو التحليل العاملي لبعض المـــتغيرات المتعلقة بالجريمة، والمستوى الاجتماعي الاقتصادي لأفراد مجتمع معين. إن العوامل المشتركة Common Factors المستخلصة من هذا التحليل تعبر عـــن الفـــروق الفردية بين المختبرين، والتفرد هنا يكون تفرد نمط الفرد في أبعاد مشتركة بـــين الأفراد.

(٢) التحليل العاملي للأفراد في سمات معينة في موقف معين، ويُطلق عليه Q-Factor Analysis: وهذا النمط من التحليل العاملي يمكن اعتباره عكس الـــنمط الأول لذلك يطلق عليه أحياناً التحليل العاملي العكسي Inverse Factor Analysis، وذلك لأن الصفوف تحل محل الأعمدة في المصفوفة، حيث يكـــون الأفـــراد أو الكينونات هي المتغيرات التي تخضع للتحليل العاملي.

(٣) التحليل العاملي للسمات التي تميز فرداً معيناً في مواقف متسلسلة، ويطلق عليـــه P-Factor Analysis: وهنا تكون الارتباطات بين مجموعة من المتغيرات لفـــرد واحد في عدد من المواقف أو الظروف الكافية،ويستخدم هذا النمط للكشف عن السمات المتفردة Unique Traits وبخاصة السمات الديناميـــة داخـــل الفـــرد،

والعوارض الظاهرة على الفرد،كما يستخدم هذا النمط في الدراسات الطولية لمجتمع معين لرصد وتحليل ظاهرة معينة على فترات زمنية طويلة. وهنا تكون العوامل التي يتم استخلاصها تقيس المكونات المستقلة للتغير الذي يحدث في سمات أحد الأفراد، أو الخصائص لكينونة معينة. وعلي الرغم من أن هذا النمط من التحليل العاملي غير شائع الاستخدام، إلا أن له فائدة كبيرة في دراسة الحالات الفردية ، حيث يكون محور الاهتمام هو الفرد، والتغيرات الاجتماعية التي تحدث داخل جماعة معينة، غير أنه ينبغي أن يكون هناك تباين متسق في الدرجات بالمتغير الذي يجري قياسه في الظروف المختلفة، كما يجب ألا تؤثر القياسات في ظرف معين علي القياسات في ظرف آخر تأثيراً له دلالة. غير أن هذا النمط من التحليل لا يناسب الخصائص أو السمات المستقرة لدى الفرد، كما لا يكون مناسبا إذا كان تباين الدرجات عبر الزمن يمكن أن يُعزى بدرجة كبيرة إلى تأثير الخبرة أو المران أو أخطاء القياس.

خامساً: طرق تدوير العوامل :

إن الكثير من طرق وأساليب التحليل العاملي تؤدي إلى نتائج يصعب تفسيرها في ضوء العوامل المستخلصة، ويرجع ذلك إلى أن مصفوفة التشبعات العاملية يتم التوصل إليها اعتباريا، ونعني بذلك أنه يمكن إعادة تكوين مصفوفة الارتباطات الأصلية عن طريق عدد لا نهائي من المحاور ، وعلى الرغم من أن مصفوفة التشبعات العاملية تلخص تباين العوامل المشتركة التي تنطوي علي مجموعة المتغيرات، إلا أننا لا نستطيع باستخدامها تفسير بني المتغيرات تفسيراً علمياً، فهذا التفسير ينبغي أن يستند إلى الشكل التكويني للمتغيرات Variables Configurations في فضاء العوامل Factor Space. ويتطلب الكشف عن هذا الشكل التكويني من خلال تدوير Rotation هذه المحاور المرجعية الاعتبارية بما يمكننا من تصور المتغيرات في فضاء يشتمل علي عدد محدد من الأبعاد. وهذا الشكل التكويني ينبغي أن يحقق متطلبات التكوين البسيط Simple Structure لتيسير التفسير العلمي للعوامل التي يتم استخلاصها.

وهناك مجموعة من المحكات التي يمكن الاستناد إليها في تدوير المحاور إلى مواقع جديدة، وذلك لتقليل عنصر الذاتية عند إجراء ذلك، وتحقيق ما يعرف بالتكوين العاملي البسيط، وهذه المحكات هي:

(١) أن يشتمل كل صف من مصفوفة العوامل علي تشبع صفري واحد، أو قريب من الصفر.

(٢) أن يشتمل كل عمود من أعمدة مصفوفة العوامل على عدد من التشبعات الـــتي تقترب من الصفر تساوي تقريبا عدد العوامل.

(٣) أن يكون لكل زوج من العوامل (الأعمدة) عدة متغيرات تشبعها بأحد العوامـــل صفر تقريباً، وتشبعها بالعوامل الأخرى كبيراً.

(٤) أن يكون لأي زوج من العوامل – إذا كان عددها أكثر مـــن أربعـــة – عـــدة متغيرات تتشبع بكل منها تشبعاً صفرياً أو قريباً من الصفر.

(٥) أن يكون عدد التشبعات غير الصفرية للمتغيرات بكل زوج من العوامل (كل من العمودين) صغيراً.

وتفيد هذه المحكات في تبسيط جميع المتغيرات لتيسير تفسير العوامل،واســتنادا إلى هـــذه المحكات يمكن تعيين الإحداثيات الجديدة لجميع النقط التي تمثل التشبعات العاملية الموضحة في الجدول السابق بالنسبة للمحاور المرجعية بعد تدويرها، حيث يوضـــح الجـــدول الآتي مصفوفة التشبعات العاملية والاشتراكيات للمتغيرات بعد التدوير :

المتغيرات	العوامل بعد التدوير		
	العامل الأول	العامل الثاني	الاشتراكيات (h^2)
١	٠,١٠	٠٨٦.	٠,٧٥
٢	٠,٠٤	٠,٧٧	٠,٥٩
٣	٠,٧٦	٠,٠٨	٠,٥٨
٤	٠,٨٣	٠,٠٣	٠,٦٩
٥	٠,٠٥	٠,٧٧	٠,٦٠
٦	٠,٩٣	٠,٠٥	٠,٨٧
٧	٠,٨٠	٠,١١	٠,٦٥
٨	٠,٠٧	٠,٨٧	٠,٧٦
التباين الذي يمكن تفسيره	٢,٧٩	٢,٧٠	٥,٤٩
نسبة التباين الكلي	٠,٣٥	٠,٣٤	٠,٦٩
نسبة التباين المشترك	٠,٥١	٠,٤٩	١,٠٠

ويتضح من هذا الجدول أن تدوير العاملين أدى إلى إعادة توزيع تشبعات المتغيرات بكـــل منهما. غير أنه يلاحظ أن مجموع مربعات التشبعات العاملية في أي صف من صـــفوف الجدول (الاشتراكيات) لم تتغير عما كانت عليه قبل التدوير،وهذا يعني أن العوامل بعـــد تدويرها تفسر نفس مقدار التباين الذي تفسره تلك العوامل قبل التدوير،كما يلاحظ أن مجموع حاصل ضرب التشبعات في أي صفين في الجدول مساو للمجموع نفـــسه قبـــل

التدوير. وهذا يعني أن التباين المشترك الذي يمكن تفسيره بواسطة العاملين لم يتغير نتيجة للتدوير العاملي، أي أن العوامل بعد التدوير تتميز بنفس خصائصها قبل التدوير، ولكن تفسيرها أصبح أكثر سهولة عما كان عليه قبل التدوير. لذلك يوصي كثير من خبراء الإحصاء تدوير العوامل للإفادة من مزايا هذا التدوير في تفسير العوامل تفسيراً يتفق ومجال الدراسة وطبيعة المتغيرات المراد تحليلها عاملياً.

وللتحقق من مدي توافر محكات التكوين البسيط التي أشرنا إليها تفحص التشبعات في الجدول،ونضع علامة (×) مكان التشبعات التي تساوي أو أكبر من ٠,٣٣ بينما نضع (صفر) مكان التشبعات التي تقل عن ذلك، أي التشبعات التي لا يعتد بها، كما هو مبين بالجدول التالي:

المتغيرات	العامل الأول	العامل الثاني
١	صفر	×
٢	صفر	×
٣	×	صفر
٤	×	صفر
٥	صفر	×
٦	×	صفر
٧	×	صفر
٨	صفر	×

ويتضح من هذا الجدول أن كل صف يشتمل علي تشبع صفري واحد، وكـل عمـود يشتمل علي أربعة تشبعات صفرية، ولا يوجد في أي صف علامة(×) للعـاملين معـا، وبذلك يسهم التدوير العاملي في تبسيط البنية العاملية لمجموعة المتغيرات.

ويعتمد التدوير العاملي البياني علي خبرة الباحث وكفاءته في اختيار المحاور المرجعية التي تحقق المحكات التي سبقت الإشارة إليها وكذلك علي حكمه السليم أثناء إجراء عمليـات التدوير. ومع التطورات التقنية المتلاحقة تحول الاتجاه إلى ابتكار طرق تحليليـة للتـدوير العاملي تستند إلى قواعد رياضية محددة يتم تنفيذها باستخدام برامج حاسـوب معياريـة جاهزة، لذلك فإن هذه الطرق لا تتطلب أحكاما مسبقة من الباحث،ولا تختلف نتائجها باختلاف مستخدميها،وتوجد طرق تحليلية متعددة لإجراء التدوير العـاملي باسـتخدام الحاسوب، ولكن لا توجد طريقة معينة تناسب جميع أنواع البيانات، فعلي الرغم مـن أن هذه الطرق يتم إجراؤها آليا بالاستناد إلى حكم الباحث، إلا أن هذا لا يمنعه من الحكـم على الطريقة التي يختارها لإجراء التدوير العاملي من حيث مد ى ملاءمتها للبيانات الـتي يتم تحليلها تحليلها، غير أن جميع هذه الطرق تستخدم علم التفاضل وجبر المـصفوفات في تعظيم Maximize أو تصغيرMinimize دالة معينة تتعلق بالتشبعات العاملية، وتـستند معظم طرق تدوير العوامل إلى مبدأ التكوين البـسيط Simple Structure للمـصفوفة العاملية ، وهناك طرق تحليلية للتدوير العاملي المتعامد Orthogonal Rotation وطرق أخرى للتدوير المائل Oblique Rotation ، وهذه الطرق ضمن حزم البرامج الإحصائية للحاسوب، وفيما يلي نبذة مختصرة عن تلك الطرق:

١- الطريقة Varimax: وتُعد أكثر الطرق استخداماً لإجراء التدوير العاملي المتعامـد، وتؤدي إلى نتائج جيدة، والكثير من الحزم الإحصائية الجاهزة تتضمن هذه الطريقة الـتي اقترحها كايزر Kaiser عام ١٩٥٨. وقد حاول عن طريقها تبسيط الصفوف والأعمدة في المصفوفة العاملية للتوصل إلى التكوين البسيط،وتعتمد هذه الطريقة علي تدوير جميـع العوامل Factors الممكنة مثنى مثنى علي حدة إلى أن يتم تعظـيم Maximizing دالـة معينة تتعلق بالتشبعات العاملية، أي أن هذه الطريقة يتم إجراؤها في خطوات متتالية مـن خوارزمية معينة Algorithm تحاول تعظيم مجموع تباينات مربعات التشبعات في أعمـدة

المصفوفة العاملية، وهذا يحقق أحد محكات التكوين البسيط ألا وهو وجود بعض التشبعات المرتفعة وبعض التشبعات الصفرية في كل عمود من أعمدة هذه المصفوفة، كما أن تربيع قيم التشبعات يجعل تلك القيم في الأعمدة موجبة. وقبل إجراء هذه التربيعات يتم تربيع قيم تشبعات كل صف من صفوف المصفوفة العاملية، وقسمة مجموع التربيعات لكل صف علي المجموع الكلي للمربعات، وذلك لإحداث تكافؤ في أهمية المتغيرات في تحديد نتائج التدوير. وقد أوضحت الدراسات المختلفة فاعلية هذه الطريقة وعدم اختلاف نتائجها عن النتائج من الطرق الأخرى.

٢- **الطريقة** Quartimax: تُعد من الطرق التحليلية المبكرة لإجراء التدوير المتعامد، وتهدف هذه الطريقة إلى جعل تباين التشبعات في كل صف من صفوف المصفوفة العاملية أكبر ما يمكن، وهنا يعمل تدوير العوامل على جعل المتغير مشبعاً بعامل واحد لكي يكون مقياساً نقياً لهذا العامل، كما أن تعظيم مجموع التباينات في كل صف من هذه المصفوفة يحقق أحد محكات التكوين البسيط، ولتلافي الإشارات السالبة لقيم التشبعات العاملية يتم تربيع هذه التشبعات، وبذلك تهدف هذه الطريقة لتعظيم مجموع تباينات صفوف المصفوفة العاملية. وعلى الرغم من أن هذه الطريقة تستخدم في بعض البحوث، إلا أنها ثبت عدم فاعليتها في التوصل إلى التكوين البسيط أو العوامل المشتركة، حيث إنها تؤدي في كثير من الأحيان إلى عامل عام بعد التدوير، لذلك يفضل استخدام الطريقة السابقة Varimax.

٣- **الطريقة** Promax : وهي تستخدم في التدوير المائل Oblique Rotation للمصفوفة العاملية، وتضمها أيضا الحزم الإحصائية للحاسوب. ويلجأ بعض الباحثين إلى استخدام التدوير المائل الذي يؤدي إلى استخلاص عوامل مترابطة، أي ليست متعامدة كما في التدوير المتعامد، حيث إن المحاور المرجعية للعوامل بينها تحصر زوايا ليست قائمة، ويرجع ذلك إلى أن هؤلاء الباحثين ربما يودون مقارنة نتائج التدوير المتعامد بنتائج التدوير المائل للتوصل إلى استنتاجات معينة، أو ربما يودون تحديد وتفسير أبسط تكوين ممكن للبنية العاملية والتوصل إلى علاقات عليةCasual، وفي هذه الطريقة توضع قيم الاشتراكات في الخلايا القطرية لمصفوفة الارتباطات بدلاً من الواحد الصحيح، وإجراء التدوير المائل علي المصفوفة العاملية الناتجة.

سادساً: تقدير درجات العوامل

يتم استخلاص العوامل من مصفوفة الارتباطات بين المتغيرات،وبخاصة بعد تدوير المصفوفة العاملية بحيث تحقق محكات التكوين البسيط قدر الإمكان. غير أن التحليل العاملي لا يقف عند هذا الحد، إذ ربما يود الباحث التوصل إلى نمط أو بروفيل في ضوء العوامل أو الأبعاد أو التكوينات الفرضية المستخلصة. فالأفراد أو الكينونات يمكن تحديد أو تقدير وقع كل منها بالنسبة للأبعاد أو العوامل بأسلوب منهجي يماثل أسلوب تحديد موقعها بالنسبة للمتغيرات الأصلية، وقد يريد الباحث إيجاد الارتباط بين متغيرات أخرى – لم تشملها مصفوفة الارتباطات الأصلية– وبين العوامل المستخلصة من هذه المصفوفة لبناء موازين Scales أو مؤشرات Indices أو لإجراء دراسات تالية، غير أن درجات العوامل لا نستطيع حسابها بطريق مباشر فيما عدا بعض الحالات المتعلقة بالمكونات الأساسية Principal Components ، أو التحليل العاملي لصور المتغيرات، وإنما ينبغي تقديرها إحصائياً ولعل أبسط طرق تقدير درجات العوامل Factor Scores هي أن نفصل المتغيرات التي يتبين أن تشبعاتها العاملية بعاملين تساوى أو أكبر من قيمة معينة ولتكن (٠,٥٠) ومن ثم نجمع الدرجات الخام لهذه المتغيرات لنحصل عل قيمة تقديرية لدرجة العامل الخاصة بفرد معين. أما المتغيرات التي تشبعاتها سالبة وتساوى (– ٠,٥٠) أو أقل فإن درجاتها الخام تطرح بدلاً من أن تجمع، وذلك لأنها ترتبط بالعامل ارتباطاً سالباً. ونظراً لأن هذه الطريقة تعتمد على الدرجة الفاصلة في اختيار المتغيرات فإنه يفضل تحويل جميع درجات المتغيرات التي تفوق هذه الدرجة إلى درجات متحدة في المتوسط والانحراف المعياري قبل إجراء عمليتي الجمع أو الطرح، إذ إن ذلك يضمن أن جميع المتغيرات تكون متساوية الوزن في تحديد درجات العوامل. كما يفضل في بعض الأحيان إعطاء أوزان متباينة لدرجات المتغيرات بعد توحيد متوسطاتها وانحرافاتها المعيارية، وهذه الأوزان تكون مساوية لتشبعات المتغيرات بالعامل المعين،وتتميز هذه الطريقة الأخيرة بأنها تسمح للمتغيرات التي حازت على أعلى التشبعات بالعامل– بالتأثير الأكبر في تقدير درجات العوامل.

وتعتمد بعض الطرق الإحصائية الأخرى المستخدمة في تقدير درجـات العوامـل علـى معادلات الانحدار المتعدد، فالعوامل المستخلصة من التحليل العاملي يمكـن اسـتخدامها كمتغيرات في معادلة انحدار، وتستخدم الدرجات المعيارية للمتغيرات في التنبؤ بـدرجات العوامل كمتغير محك والحصول على قيم معاملات الانحدار المعيارية. وهذه العوامل قـد تكون مرتبطة أو غير مرتبطة، أي تم استخلاصها بالتدوير المائل أو المتعامد، غير أنه يفضل استخدام عوامل متعامدة حيث إنه يمكن تفسير نتائجها تفسيراً مباشراً وواضـحاً، ومـن الجدير بالذكر أن برامج الحاسوب تتضمن إجراء تحليل الانحدار المتعدد والحصول علـى درجات العوامل أفراد العينة.

سابعاً: تسمية العوامل وتفسيرها:

بعد استخلاص العوامل وتدويرها تدويراً متعامـداً أو مـائلاً، والتوصـل إلى مـصفوفة التشبعات العاملية بعد التدوير، ينبغي تفسير البنية العاملية الناتجة عن التحليـل، وإضـفاء معنى على العوامل بحسب مجال الدراسة، وذلك بتسمية العوامل تسمية مناسبة، والاستعانة بالتمثيل البياني لتوضيحها واستيعابها. وتعتمد تسمية العوامل Naming Factors علـى الهدف من التحليل ومنظور الباحث بما يتعلق بهذه العوامل،فقد يهدف الباحـث- مـن استخدام التحليل العاملي – توضيح مفاهيم مفيدة في مجال يفتقر إلى المفاهيم الواضـحة، وبعض الباحثين يهدف إلى التوصل إلى علاقات علية تفسر أسباب ظاهرة معينـة، ففـي الحالة الأولى تصف العوامل Factor العلاقات القائمة بين المتغيرات، وتصنفها في ضـوء التسمية التي يقترحها الباحث، أما في الحالة الثانية فإن الباحث ينظر إلى العوامل كأسباب تنطوي عليها هذه العلاقات، وتكون التسمية متعلقة بهذه الأسباب، وفي بعض الحـالات الأخرى ربما لا يحاول الباحث تسمية العوامل أو إضفاء معنى معين عليها، وإنما يرمز لهذه العوامل برموز معينة، ويوضح فقط تشبعاتها ببعض العوامل.

ومما لا شك فيه أن التسمية المناسبة للعوامل تفيد في التواصل بين الباحثين، وتعـبر عـن المضمون الوصفي لهذه العوامل أو فحواها مما ييسر فهم معانيها، ولعلنا نلاحـظ مـدى غموض وتشابك كثير من التسميات والمـصطلحات والمفـاهيم في العلـوم الـسلوكية

والاجتماعية، مما يؤدى إلى إرباك ملحوظ لدى الباحثين الذين يستخدمون تلك المصطلحات،كما أن تسمية العوامل قد يفيد في اقتراح فروض تتطلب مزيداً من البحث والتجريب والتحقق.

وينبغي أن تكون التسمية بسيطة ومختصرة ومعبرة عن المفهوم المعين، كما ينبغي اختيار التسمية المناسبة للعوامل بحيث ترتبط بالهدف الأساسي من إجراء التحليل العاملي،فإذا كان الهدف وصف العلاقات المتعددة بين المتغيرات، فإن تسمية العوامل يمكن أن تكون وصفية بحيث تعكس مضمون المتغيرات التي تشبعاتها بعامل معين مرتفعة وتشبعاتها بالعوامل الأخرى قريبة من الصفر. أما إذا كان الهدف هو الكشف عن العلاقات العلية، فإن أنماط العلاقات الإمبيريقية بين المتغيرات يفرض أنها تعكس الآثار المشتركة للمتغيرات، ويؤكد ذلك وجود تجمعات من المتغيرات المرتبطة فيما بينها ارتباطاً مرتفعاً مما يدل على أن هناك عاملاً مشتركاً ينطوي عليها، كما أن التحليل العاملي ربما يكشف عن علاقات علية. وهناك عدة اعتبارات ينبغي مراعاتها في تسمية العوامل الوصفية أو العلية، ومن أهم تلك الاعتبارات ما يلي:

١-عند تفسير العوامل ينبغي الأخذ بعين الاعتبار المتغيرات غير المرتبطة بأحد العوامل، أي التي تكون تشبعاتها بهذا العامل صفرية أو قريبة من الصفر، فالتسمية ينبغي أن تعكس ما يتضمنه العامل وما لا يتضمنه.

٢- تدل مربعات التشبعات العاملية على تباين المتغير الذي يمكن تفسيره بمعلومية عامل متعامد، مما يساعد في تحديد الوزن النسبي الذي ينبغي أن يعطى للمتغير في تفسير العامل.

٣-بعض العوامل المحدودة أو الصغيرة ربما يصعب تفسيرها نتيجة الأخطاء العشوائية الناجمة عن العينة،ففي هذه الحالة لا يجب تقسيم هذه العوامل، وإنما يمكن أن يتناولها عدد من الدراسات المتكررة للتحقق مما إذا كانت متسقة أو غير متسقة.

٤-يمكن التحقق من قيم التشبعات العاملية التي لا تعزى إلى الصدفة باستخدام اختبارات الدلالة الإحصائية.

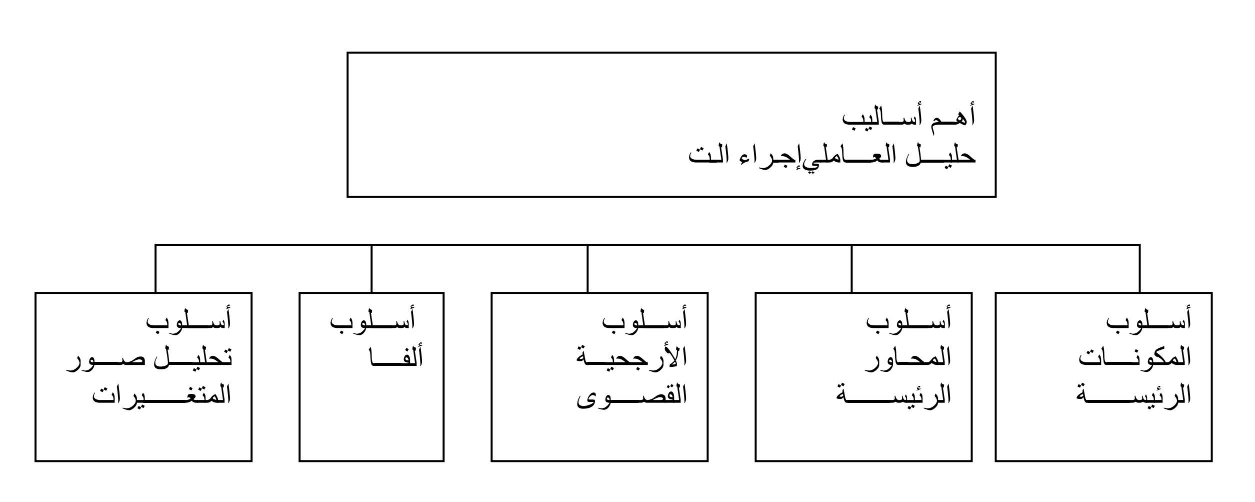
٥-إن تسمية العوامل التي بعض تشبعاها موجبة مرتفعة، وبعضها الآخر سالبة مرتفعة، ينبغي أن تعكس هذه الثنائية القطبية Bipolarity، مثل (مرتفع- منخفض)، أو (حضري – ريفي).

٦-بعض العوامل التي يتعذر تسميتها يمكن تفسيرها بدرجة أفضل بأن نعكس إشارة بعض قيم تشبعاها لأنه عندما نعكس إشارة أحد المتغيرات فإن هذا يؤدى بالتالي إلى أن تنعكس إشارة مقياس أو ميزان المتغير.

٧-تسمية المتغيرات بصفات تشير إلى اتجاه العلاقة بالعوامل ييسر تفسير هذه العوامل (مثل انخفاض الوفيات، وارتفاع الدخل القومي).

ثامناً: أساليب التحليل العاملي:

لقد أسهمت التطورات التقنية إسهاماً كبيراً في تيسير عملية تحليل البيانات متعددة المتغيرات مما حث العلماء على تطوير نماذج إحصائية متقدمة وأساليب أكثر دقة لاستخلاص العوامل من مصفوفات الارتباطات، وأصبح من الممكن الآن إجراء التحليل العاملي لمصفوفات تشتمل على متغيرات كثيرة في وقت قصير وبدقة عالية من خلال الحاسوب، ويوضح الشكل التالي بعض أهم الأساليب المعاصرة لإجراء التحليل العاملي:



وفيما يلي توضيح موجز لتلك الأساليب:

١- **أسلوب المكونات الرئيسية** Principal Component Procedure:

تشتمل كثير من بيانات البحوث على قياسات لمتغيرات تكون مرتبطة فيما بينها، وفي بعض الأحيان يريد الباحث إجراء تحويل على هذه البيانات بحيث تصبح المتغيرات غــير مرتبطة، كأن يريد معرفة عدد مصادر التباين المستقلة استقلالاً خطياً في النظام الذي تمثله هذه المتغيرات، فالعلاقات القائمة بين عدد من المتغيرات غير المرتبطة يمكن تلخيصها بعامة في عدد أقل من المكونات الرئيسة Principal Components لــذلك فــإن أســلوب المكونات الرئيسة يعد أسلوباً لتكوين متغيرات الملاحظة، ولكي نحصل علــى مجموعــة وحيدة من المعاملات يتم تعريف المكونة الرئيسة الأولى بأنها التركيب الخطى من المتغيرات التي يكون تباينها أكبر من تباين جميع الدوال الخطية التي تشتق من مجموعة المتغيرات مجال البحث. أما المكونة الرئيسة الثانية فإنها أيضا تعد تركيباً خطياً من المتغيرات التي تباينــها أكبر من جميع الدوال الخطية لمجموعه المتغيرات المتعامدة على المكونــة الرئيــسة الأولى، وهكذا في المكونات الرئيسة التالية حتى ينتهي تحليل التباين الكلى. ويطلق على معاملات المكونات الرئيسة "تشبعات المكونات الرئيــسة Principal Component Loadings" . أما قياسات المكونات الرئيسية لكل فرد فتسمى "درجات المكونة الرئيــسة Principal Component Scores" وكل من هذه الدرجات عبارة عن مجموع ما تسهم بــه جميــع تشبعات المتغيرات في هذه المكونة الرئيسة. وينبغي ملاحظة أن جميع المكونات الرئيــسة تكون مستقلة بعضها عن البعض الآخر، أي غير مرتبطة، وكل منها يسهم بأكبر قدر في التباين المتبقي من المكونة الرئيسة السابقة عليها، كما أن مجموع تباين المكونات الرئيــسة جميعها يساوى مجموع تباين المتغيرات الأصلية.

٢–أسلوب المحاور الرئيسة Principal Axes Method:

يعد أسلوب المحاور الرئيسة الذي قام بتطويره هوتيلنج Hotelling من أكثر أساليب التحليل العاملي المعاصرة استخداماً في استخلاص العوامــل مــن المــصفوفة الكاملــة

للارتباطات بين المتغيرات ، ويختلف هذا الأسلوب عن أسلوب المكونات الرئيسة الـــذي أوضحناه في أن الخلايا القطرية لمصفوفة الارتباطـــات يوضـــع فيهـــا القـــيم التقديريـــة للاشتراكيات (أي مجموع مربعات تشبعات المتغير بجميع العوامل)، بينمـــا في أســـلوب المكونات الرئيسة يوضع الواحد الصحيح في كل من هذه الخلايا. ويهتم أسلوب المحـــاور الرئيسة باستخلاص عدد معين من العوامل التي تفسر أكبر قدر ممكـــن مـــن التبـــاين في مصفوفة الارتباطات، ويعتمد هذا الأسلوب على مبادئ أساسية في هندســـة المتجهـــات وجبر المصفوفات، وبخاصة تحليل المتجهات لمعاملات الارتباط، والجذور الكامنة Latent Roots. وأهم ما يميز هذا الأسلوب أن كل عامل من العوامل المستخلصة يفسر أكبر قدر ممكن من تباين المتغيرات المراد تحليلها عاملياً، فالعامل الأول الذي يستخلص من مصفوفة الارتباطات يشتمل على التركيب الموزون لجميع المتغيرات التي تـــؤدى إلى أعلـــى قـــيم لمربعات الارتباطات بين هذه المتغيرات وهذا العامل، وذلك لأن هذه القيم تعـــد بمثابـــة التباين الذي يفسر هذه الارتباطات،وهذا يعنى أن مجموع مربعات العمـــود الأول لبنيـــة العامل سوف تكون أكبر ما يمكن، أما العامل الثاني فيتم استخلاصه بحيـــث لا يكـــون مرتبطاً بالعامل الأول، ويفسر أكبر قدر ممكن من تباين مصفوفة البواقي بعـــد اســـتبعاد العامل الأول، وهكذا يتم استخلاص بقية العوامل، وبذلك تفـــسر مجموعـــة العوامـــل المستخلصة أكبر قدر ممكن من التباين.

كما أن هذا الأسلوب يتميز بأن مصفوفة الارتباطات التي يعاد تكوينـــها Reproduced Correlation Matrix تعد بمثابة أفضل تقدير للمربعات الصغرى لمصفوفة الارتباطـــات الأصلية التي تكون قيم خلاياها القطرية الواحد الصحيح أو القيم التقديرية للتشبعات، أي أن مجموع مربعات الفروق بين المصفوفة الأصلية والمصفوفة الناتجة تكون أقل ما يمكـــن. وهذا يعنى أن أسلوب المحاور الرئيسة يؤدى إلى استخلاص أقل عدد من الأبعاد المطلوبـــة لإعادة تكوين أو تفسير البيانات الأصلية خطياً، ويمكن استخلاص كل من العوامل علـــى حدة أو استخلاص جميع العوامل في آن واحد. والعمليات الحسابية التي يتطلبـــها هـــذا الأسلوب تبدأ عادة بقيم مبدئية، ثم يجرى تحويل هذه القيم عن طريق التكـــرار المتتـــالي Iteration وذلك برفع مصفوفة الارتباطات إلى قوى (أسس) متتابعة Successive لحين

التوصل إلى قيم مستقرة في حدود معينة، ومن هذه القيم يتم تحديد أول جــذر كــامن Eigen Value، وأول متجه كامن Eigen Vector، ومنها نحسب التــشبعات العامليــة Factor Loadings، ثم تحسب قيم خلايا مصفوفة الارتباطات بعــد إعــادة تكوينــها Reproduced Correlation Matrix وتطرح من مصفوفة الارتباطات الأصلية، وتكرر هذه العمليات بالنسبة لبواقي مصفوفة الارتباطات، ومع هذا التكرار لاستخلاص العوامل تصبح مصفوفات البواقي أقل فأقل إلى أن تصبح البواقي صغيرة جداً بما يمكن عــزوه إلى الصدفة، أو إلى تباين غير ذي دلالة،ويرجع ذلك إلى أن إسهامات العوامــل المتتاليــة في التباين تقل تدريجياً، فالعامل الأول يقيس أكبر قدر من التباين، وتقل الإسهامات النــسبية للعوامل التالية على التوالي، ولتفسير العوامل المستخلصة يجرى تدويرها تــدويراً متعامــداً

٣-أسلوب الأرجحية القصوى Maximum Likelihood method:

يعتمد أسلوب المحاور الرئيسة وغيره من الأساليب التي أوضحناها على اختزال مــصفوفة بواقي الارتباطات بين المتغيرات، ولكن أسلوب الأرجحية القــصوى (MLM) تتركــز إجراءاته الإحصائية في التوصل إلى أفضل قيم تقديرية للتشبعات العاملية، وقد تم تطــوير هذا الأسلوب من أساليب التحليل العاملي بحيث يسمح باختيار الدلالات الإحصائية لعدد العوامل التي يمكن استخلاصها من المصفوفة الإمبيريقية لمعاملات الارتباط، ويعمل علــى استخلاص مجموعة من العوامل المشتركة، وعامل خاص بكل متغير.

ويفيد هذا الأسلوب في التغلب على المشكلة التي تتعلق بالتمييز بين تبــاين العوامــل المشتركة (الاشتراكيات) والتباين الخاص لأن هذا يتطلب تقدير الاشتراكيات، فارتبــاط كل متغير بنفسه Self – Correlation في خلايا مصفوفة الارتباطات ينبغي أن تكون هي نفسها قيم الاشتراكيات للمتغير،والأساليب الشائعة الاستخدام في التعامــل مــع تلــك المشكلة تتمثل في البدء بقيم تقديرية مبدئية للاشتراكيات وربما تعدل هذه القــيم أثنــاء التحليل، وباستخدام هذه القيم نحاول استخلاص العوامل بحيث تكون هناك مطابقة جيدة بالبيانات الملاحظة، لكن أسلوب الأرجحية القصوى لا يتطلب هذه القيم التقديرية المبدئية مقدماً،إذ إن الاشتراكيات تعتمد على افتراض عدد العوامل المشتركة.

كما أن أسلوب الأرجحية القصوى يختلف عن غيره من الأساليب في تعامله مع مصفوفة الارتباطات باعتبارها تشتمل على عينة من المتغيرات، فالأساليب الأخرى تفترض أن هذه المصفوفة تؤدى إلى نتائج تنطبق عل مجتمع المتغيرات الذي استمدت منه هذه العينة دون استخدام الأساليب الإحصائية للتحقق من ذلك بينما يراعى أسلوب الأرجحية القصوى الفروق بين الارتباطات القائمة بين التغيرات الملاحظة والقيم النظرية في مجتمع المتغيرات الذي يتم معاينته،ويؤدي هذا الأسلوب إلى تقديرات Estimates لمجتمع أوزان العوامل Universe Factor Weights في إطار النموذج العاملي المتعلق به

من هنا فإن أسلوب الأرجحية القصوى يتطلب فرضية صفرية تتعلق بعدد العوامل المشتركة التي يمن استخلاصها من مصفوفة الارتباطات، وفي هذه الفرضية تستخلص العوامل واشتراكياتها، وهذه الفرضية الصفرية مؤداها أن التباين الكلى في مجتمع المتغيرات أمكن استخلاصه بواسطة عدد افتراضي من العوامل.

وعلى الرغم من أن أسلوب الأرجحية القصوى يعد أسلوباً جيداً للتوصل إلى عوامل مشتركة بدلاً من المكونات الأساسية، كما يزودنا باختبار الدلالة الإحصائية فيما يتعلق باستخلاص كل من العوامل المشتركة المتتالية–إلا أنه يعاب عليه تعقد اشتقاقاته الرياضية وكثرة عمليات الجبرية والحسابية وتحليلاته العددية اللازمة للتعامل مع المعادلات المعقدة الناتجة عن إجراءاته. وقد انتشر استخدام الحاسوب ذي السرعة الفائقة والبرمجيات المتقدمة التي تم تطويرها لتنفيذ الإجراءات والعمليات التي يتطلبها هذا الأسلوب، إلا أنه لا تزال هذه الإجراءات معقدة، وتتطلب قدراً كبيراً من ذاكرة الحاسوب، ووقتاً غير قصير لإيجاد الحل العاملي، وذلك بالتكرار المتتالي Iteration للتوصل إلى التقارب Convergence الذي ربما يصعب تحققه في بعض الحالات. ومن هذا يتبين أن أسلوب الأرجحية القصوى يمكن اعتباره أسلوباً كشفياً لاختزال المتغيرات، كما يمكن اعتباره أسلوباً تأكيدياً للتحقق من صحة الفروض.

٤-أسلوب ألفا Alpha Factor Analysis:

الهدف من هذا الأسلوب تقدير العوامل المشتركة من مجموعة من المتغيرات التي يتم انتقاؤها بحيث يكون معامل ارتباطها بنطاق شامل Universe من العوامل المشتركة أكبر ما يمكن، ويفترض أن الأفراد الذين تقاس لديهم هذه المتغيرات المنتقاة يشكلون المجتمع المستهدف،وقد استمد هذا الأسلوب مسماه من معامل الثبات ألفا(α) ، ويمكن اعتبار معامل ألفا مربع معامل الارتباط بين عامل مشترك في مجموعة المتغيرات المنتقاة والعامل المشترك المناظر في النطاق الشامل، ويسمى عندئذ معامل إمكانية التعميم العاملي المشترك Factor Generalizability، لذلك فإن هذا الأسلوب يعد من الأساليب السيكومترية.

ويبدأ التحليل بتقدير قيم للاشتراكيات، وتستخلص العوامل الأساسية من مصفوفة يتم تكوينها وفق معادلة خاصة تعتمد على مصفوفة قطرية Diagonal Matrix توضع في خلاياها الجذور التربيعية للاشتراكيات، ثم يعاد حساب الاشتراكيات من هذه المصفوفة، وتستخدم الاشتراكيات الناتجة بديلاً عن الاشتراكيات المبدئية المقدرة،ويتم تكوين مصفوفة جديدة تتضمن هذه الاشتراكيات المحسوبة حيث تستخلص منها العوامل، وتكرر هذه العمليات على التوالي لحين التوصل إلى التقارب في الاشتراكيات Communalities، ومن ثم يعاد وزن Re- Scaling مصفوفة العوامل.

٥-أسلوب صور المتغيرات Image Factor Analysis:

ويعد هذا الأسلوب من الأساليب السيكومترية في التحليل العاملي، ويؤدي غالباً إلى نتائج قريبة إلى حد ما من نتائج أسلوب المحاور الرئيسة، ويختلف أسلوب صور المتغيرات عن الأساليب التقليدية التي تستخدم في استخلاص العوامل المشتركة، فهذه الأساليب ركزت على مشكلة تحديد مجموعة صادقة من الاشتراكيات للمتغيرات أكثر من تركيزها على العلاقات بين المتغيرات، فالعوامل المشتركة يجرى استخلاصها على التوالي إلى أن تصبح مصفوفة بواق الارتباطات غير دالة، أي يجرى استخلاص كل عامل على حدة، وإحلال مصفوفة الارتباطات الجزئية بين المتغيرات بعد عزل أثر هذه العوامل محل مصفوفة

الارتباطات الأصلية. وتنتهي العمليات عندما تصبح قيم معاملات الارتباط الجزئية المتبقية ضئيلة.

ولكن في أسلوب تحليل صور المتغيرات يكون التركيز على مربع معاملات الارتباط بين المتغيرات بدلاً من معاملات الارتباط الجزئية المنفصلة، أي أنه في هذا الأسلوب يكون "الجزء المشترك" من متغير في مصفوفة الارتباطات هو ذلك الجزء الذي يمكن التنبؤ به من بقية المتغيرات في المصفوفة باستخدام الانحدار المتعدد. فإذا اشتملت المصفوفة على تسعة متغيرات مثلاً، فإننا نستخدم المتغير الأول كمتغير تابع ونتنبأ به من المتغيرات الثمانية الأخرى باستخدام معادلة الانحدار الخطى المتعدد، ثم نستخدم المتغير الثاني كمتغير تابع ونتنبأ به من المتغيرات الثمانية الأخرى بما في ذلك المتغير الأول بالطريقة نفسها، وهكذا. وبذلك نحصل على مصفوفة درجات الصور Image Score Matrix (ISM) التي تشتمل على الدرجات التي يمكن التنبؤ بكل منها من الآخر في مصفوفة الدرجات الأصلية. ويمكن بعد ذلك إجراء تحليل المكونات Component Analysis لمصفوفة درجات الصور عن طريق تكوين مصفوفة التغاير لدرجات الصور Co variances Image Scores وليس مصفوفة ارتباطات هذه الدرجات،وبعد ذلك توضع قيم مربع معامل الارتباط في الخلايا القطرية لمصفوفة التغاير، ويتم تحليل هذه المصفوفة عامليا باستخدام أسلوب المكونات الأساسية أو الطريقة المركزية.

ويتميز أسلوب تحليل (صور المتغيرات) بأنه يمكننا من التوصل إلى حل متفرد لتحليل التباين المشترك، أي المتعلق بالجزء الذي يشترك فيه كل متغير مع بقية المتغيرات المراد دراستها. ونظرا لأن العوامل المستخلصة تعد بمثابة تركيب خطي من الدرجات في مصفوفة درجات الصور (ISM)، فإن تحليل صور المتغيرات يتميز بنفس ميزات أسلوب المكونات الرئيسة من حيث إمكانية الحصول مباشرة علي درجات الأفراد في عوامل صور المتغيرات (Image Factors) وليس استنادا إلى تقديرات لها كما هو الحال في بعض الأساليب الأخرى.

ولكن نظرا لأن هذا الأسلوب يعتمد علي تحليل مصفوفة التغاير بــدلاً مـــن مـــصفوفة الارتباطات، فإن مستخدمي هذا الأسلوب من غير الاختصاصيين يجدون صعوبة في تفسير نتائج هذا التحليل، فالتشبعات العاملية الناتجة ليست ارتباطات بين المتغيرات بالعوامـــل، وإنما تغايرات Co variances متغيرات الصور Image Variables مع تركيب خطي لهذه المتغيرات.

تعقيب على أساليب التحليل العاملي:

نتبين مما سبق أن أساليب التحليل العاملي تختلف فيما بينها من حيث كمية التباين المراد تحليله، والقيم التي توضع في الخلايا القطرية لمصفوفة معاملات الارتباط، فــبعض هــذه الأساليب يهتم باستخلاص عوامل مشتركة Common Factors تفسر التباين الكلي في المتغيرات، وأحيانا يكون عدد هذه العوامل تقريبا عدد المتغيرات كما في أسلوب المكونات الرئيسية Principal Component Method، حيث يوضع واحد صــحيح في الخلايــا القطرية لمصفوفة معاملات الارتباط، وهذا الأسلوب الذي يهدف لاستخلاص المكونات الرئيسية بعد أسلوب متمايزاً عن أساليب التحليل العاملي، لذلك يطلق عليه من المنظــور العلمي النموذج المغلق Closed Model.

أما أساليب التحليل العاملي الأخرى فإن بعضها يهتم باستخلاص عدد مـــن العوامـــل المشتركة المرتبطة، وعدد مساو من العوامل الخاصة Specific Factors غير المرتبطة، وهنا توضع قيم الاشتراكيات Communalities في الخلايا القطرية لمصفوفة معاملات الارتباط المراد تحليلها عامليا، لذلك فإن هذه الأساليب تستند إلى ما يُعرف بالنموذج الحر للتحليل العاملي العام Free General Factor Analysis Model، وذلك لأن بعـــض العوامـــل المشتركة ربما يتم استبعادها لعدم أهميتها، غير أن بعض أساليب التحليل العاملي الأخـــرى تهتم باستخلاص عدد من العوامل المشتركة المرتبطة وعدد آخر أقل من العوامل الخاصـــة استنادا إلى نموذج التحليل العاملي التقريبي Approximating Factor Analysis، وذلك لأنه يستند إلى أساليب إحصائية أو إجراءات تقريبية تُطبق علي البيانات الإمبيريقية، وكل من هذين النموذجين يُعد نموذجا مفتوحا Open Model من المنظور العلمي.

ويري البعض أن النموذج المغلق الذي ينتمي إليه أسلوب المكونات الرئيسية Principal Component Method لا يكون مقبولاً من منظور البحث العلمي، وذلك لأنه ليس من الممكن أن ينطوي عدد معين من المتغيرات علي عدد مماثل من العوامل التي تفسر التباين في هذه المتغيرات،إذ إن معني ذلك أن هذه المتغيرات تفسر داخليا كمجموعة فرعية منعزلة عن النطاق الشامل لهذه المتغيرات.

أما أسلوب المحاور الرئيسية Principal Axis Method فيعتمد علي مصفوفة الارتباطات المختزلة Reduced Matrix مع وضع قيم الاشتراكيات بدلا من الواحد الصحيح في خلايا المصفوفة، لذلك فإنه يعتمد علي إجراءات تقريبية Approximate، فالعمليات الجبرية المستخدمة في هذا الأسلوب تُعطي مجموعا موزونا لكل عمود من أعمدة المصفوفة بعمليات تقريب متتالية علي العكس من الطريقة المركزية Centroid Method التي تقتصر علي الجمع البسيط لقيم معاملات الارتباط في كل من أعمدة المصفوفة للحصول علي متوسط ارتباط المتغير بجميع المتغيرات الأخرى مقسوماً على الارتباط المركزي Centroid Correlation.

واعتماد أسلوب المحاور الرئيسية علي الجمع الموزون يجعل تنفيذ إجراءات هذا الأسلوب باستخدام الحاسوب أكثر تعقيداً ويتطلب وقتاً أطول، وعلي الرغم من ذلك فإن أسلوب المحاور الرئيسية يعد أكثر الأساليب استخداماً، خاصة وأن هناك برامج حاسوب فاعلة لاستخلاص العوامل بهذا الأسلوب ، أما الطريقة المركزية فلا تزال تعد من الطرق التقريبية الجيدة التي تناسب الباحث الذي لا يكون الحاسوب متاحا لديه، حيث يمكنه إجراء الخطوات المتتالية لهذه الطريقة باستخدام الآلة الحاسبة، وبخاصة إذا كانت مصفوفة الارتباطات صغيرة.

وعلى الرغم من أن أسلوب الأرجحية القصوى يتميز بالدقة الكبيرة، إلا أن إجراءات استخلاص العوامل يستغرق وقتا مكلفاً، كما أن الكثير من برامج الحاسوب التي تجري ذلك الأسلوب لا تستوعب إلا عدداً قليلاً من المتغيرات، ويتأثر عدد العوامل المستخلصة تأثراً ملحوظا بتغير عدد أفراد العينة.

أما الأسلوب ألفا Alpha Analysis، وأسلوب التحليل العاملي لصور المتغيرات Image Factor Analysis فيُعدان من الأساليب التي لها ميزات سيكومترية أكثر من ملاءمتها

لنموذج علمي معين، فالغرض الإحصائي للأسلوب ألفا رفع قيم معـــاملات التجـــانس لتقديرات درجات العوامل Factors. ولعل هذا يماثل معامل ألفا لكرونباك Cronbach's Alpha الذي يُستخدم لتقدير تجانس مفردات اختبار أو مقياس معين. أما أسلوب التحليل العاملي لصور المتغيرات فيعد بديلا عن أسلوب المكونات الرئيسية، إذ إنه يكشف عـــن المكونات التي تشترك فيها مجموعة من المتغيرات، وهذا الأسلوب يستخلص العوامل فقط من جزء مشترك للتباين المتداخل إحصائياً في مصفوفة البيانات، ،وإن كان يماثل أســـلوب تحليل المكونات الرئيسية، حيث تُعد العوامل تركيبات خطية فعليـــة لـــدرجات صـــور المتغيرات، كما أن هذين الأسلوبين غير شائعي الاستخدام في التحليل العاملي لمصفوفات ارتباطات المتغيرات وذلك لعدم ألفة كثير من الباحثين بهما، وما يتطلبانه من رياضـــيات عالية.

غير أنه ينبغي أن نوجه نظر الباحث إلى أن تحليل مجموعة واحدة من البيانات باســـتخدام أساليب مختلفة للتحليل العاملي ربما يؤدي إلى نتائج مختلفة اختلافـــاً ملحوظـــاً؛ وذلـــك لاختلاف الافتراضات التي يستند إليها كل من هذه الأساليب، وربما كان هذا أحد أوجه النقد الرئيسية التي توجه للتحليل العاملي، ولكن هذه النتائج المختلفة تُعد تفسيراً لمصفوفة الارتباطات الأصلية، وربما تكون هذه التفسيرات جميعاً صحيحة من الوجهة الرياضـــية، فالتحليل الرياضي يوضح فقط أن مجموعة البيانات نفسها يمكن تفسيرها بطرق مختلفـــة. ومما لا شك فيه أن بعض هذه التفسيرات يكون أكثر فائدة من غيره من الوجهة العلمية، فالتحليل العاملي أسلوب إحصائي رياضي لا يتضمن شيئاً يوضح أي تفسيرات أكثر فائدة من غيرها، لذلك ينبغي علي الباحث أن يستخدم أساليب أخرى إضافية تختلـــف عـــن التحليل العاملي الذي استخدمه للتحقيق من القيمة العلمية لنتائجه.

أهم محكات اختيار أسلوب التحليل العاملي المناسب:

في ضوء تعدد أساليب التحليل العاملي التي تستخدم في التوصل إلى قيم عددية للعوامل التي يتم استخلاصها من مصفوفة البيانات، يكون من المناسب أن نوضح فيما يلـــي بعـــض المحكات التي يمكن أن يسترشد بها الباحث في اختياره للأسلوب الذي يناسب دراسته:

(١) **متطلبات تجميع المتغيرات**: بعض أساليب التحليل العاملي، مثل الطريقة المركزية تتطلب تجميعاً مبدئياً للمتغيرات المرتبطة، وربما يتم ذلك عن طريق الفحص العيني للمصفوفة مما يدخل عنصر الذاتية في عملية التحليل، ولكن يقلل الزمن اللازم لإجرائها يدويا أو باستخدام آلة حاسبة، غير أن استخدام الحاسوب في التحليل العاملي أدى إلى تيسير إجراء العمليات الحسابية وقلل من عنصر الذاتية.

(٢) **عدد العوامل**: يمكن تحديد أسلوب التحليل العاملي المناسب في ضوء عدد العوامل ونوعها. فإذا كان الباحث يود افتراض أن هناك عاملاً واحداً مشتركاً وبعض العوامل الخاصة أو الطائفية، فإنه يمكنه استخدام أسلوب العاملين، غير أن بعض أساليب التحليل العاملي تحدد عدد العوامل في البيانات، والحقيقة أن تلك الأساليب تنقسم إلى قسمين، أحدهما يتعلق بتحليل مصفوفة الارتباطات كاملة دفعة واحدة، مثل أسلوب المحاور الرئيسية، وطريقة التحليل العاملي المتعدد، وتؤدي إلى عامل مشترك واحد أو أكثر، أما القسم الثاني فيتعلق بتقسيم المصفوفة إلى مجموعات فرعية للحصول علي عوامل طائفية محددة وغير متداخلة.

(٣) **مدى تعقد المتغيرات**: والمقصود بذلك عدد العوامل التي تكون قيم تشبعات متغير معين بها كبيرة، إن تحديد هذه القيم يعتمد علي تقييم الخطأ في البيانات، والارتباطات بين المتغيرات بعامة في المصفوفة، ونتائج الدراسات السابقة، كما يعتمد على ما إذا كان الباحث يود التوصل إلى جميع العوامل المتعلقة بمتغير معين، أو يختار أسلوباً يؤدي إلى عامل واحد أو عاملين.

(٤) **التعقيد العاملي**: المقصود بذلك عدد المتغيرات التي تكون تشبعاتها متوسطة، فالعوامل العامة General Factors تكون تشبعات جميع المتغيرات بها متوسطة أو مرتفعة لأنها تكشف عن النمط العام للعلاقات بين المتغيرات. أما العوامل الطائفية Group Factors فتكون تشبعات مجموعة أو مجموعتين فرعيتين فقط من المتغيرات بها متوسطة أو مرتفعة. أما العوامل الخاصة Specific Factors فيكون تشبع كل منها بعامل واحد مرتفعاً. وتتباين أساليب التحليل العاملي في درجة التعقيد العاملي الذي تتوصل إليه، لذلك ينبغي

علي الباحث أن يحدد مسبقاً ما إذا كان الأسلوب المستخدم سوف يكشف درجة التعقد العاملي المناسب للبيانات أو للغرض البحثي.

(٥)**إسهام العامل في التباين:** إن إسهام العامل في التباين الكلي للمتغير يتحدد بواسطة مجموع مربعات التشبعات العاملية لهذا المتغير بشرط أن تكون العوامل متعامدة (غير مرتبطة). وبعض أساليب التحليل العاملي يستخلص العوامل واحداً تلو الآخر بحيث يسهم كل عامل تال في التباين الكلي بنسبة أقل مما قبله، أما البعض الآخر من أساليب التحليل العاملي فيستخلص العوامل بحيث تتساوى أهميتها جميعا من حيث الإسهام في التباين الكلي. إن اختيار الباحث الأسلوب المناسب يعتمد علي طبيعة ومستوى العوامل التي يكون لها فائدة علمية.

(٦) **تأثير عوامل الخطأ:** تفترض بعض الأساليب أن قياس المتغيرات غير مشوب بعوامل الخطأ، وأن الاشتراكيات التي توضع في الخلايا القطرية لمصفوفة الارتباطات دقيقة نسبياً، ولعل الطريقة المركزية تكون مناسبة في مثل هذه الحالة.

(٧) استقرار العوامل: تؤدي بعض أساليب التحليل العاملي إلى عوامل غير محددة من الناحية الإحصائية مثل الطريقة المركزية، في حين أن أساليب المحاور الرئيسية يؤدي إلى حل متفرد لمصفوفة معاملات الارتباط، أي عوامل محددة أو مستقرة لا تختلف من تحليل إلى آخر.

(٨)التحليل الكشفي والتحليل التأكيدي: إذا كان هدف الباحث يقتصر علي الكشف عن العوامل التي تفسر مصفوفة الارتباطات بين المتغيرات، فإن أسلوب المحاور الرئيسة أو الطريقة المركزية يكون مناسباً، أما إذ كان الهدف التحقق من فرضيات معينة أو التحقق من العوامل التي استخلصتها دراسات أخرى في مجال معين، فإن أسلوب الأرجحية القصوى يكون هو الاختيار المناسب.

تاسعاً: تصميم وإجراء الدراسات العاملية:

يتطلب إجراء الدراسات العاملية تصميما ينبثق من أهداف الدراسة وما تستند إليه إطـــار نظري، ويمر إجراء الدراسة العاملية بخطوات متتالية سعيا للتوصل إلى النتائج وتفـــسيرها، وإن كانت الخطوات الأساسية في تصميم وإجراء الدراسات العاملية تتمثل في الآتي:

(١) **تحديد الهدف من الدراسة العاملية**: إن الدراسات العاملية يمكن أن تكون كشفية Exploratory أو تأكيدية Confirmatory.لذلك ينبغي علي الباحث تحديد مـــا إذا كان الهدف من دراسته الكشف عن العوامل التي تنطوي عليها مجموعة مـــن المتغيرات، أي اختزال عدد المتغيرات إلى عدد أقل من العوامل التي تعـــد بمثابـــة تركيب خطي من هذه المتغيرات من أجل وصف الظاهرة وصفاً بسيطاً مقتـــصداً في ضوء عدد محدود من المفاهيم أو التكوينات الفرضية، فعندئذ يمكنه اســـتخدام نمط التحليل العاملي الكشفي. أما إذا كان الباحث يهدف من دراسته التحقق من نظرية معينة أو اختبار صحة فرضية تتعلق بوجود عامل أو عوامـــل معينـــة، أو التحقق من وجود نمط معين من العلاقات في مجموعة من البيانات، فإنه يمكنه في هذه الحالات استخدام نمط التحليل العاملي التأكيدي.

(٢) **صياغة الفرضية أو التساؤل البحثي المتعلق بالتحليل العاملي**: وتعد هذه مـــن الخطوات المهمة التي توجه الباحث إلى التصميم المناسب لدراسته، كما ترشـــده في اختيار نموذج التحليل العاملي وأسلوبه وطريقة التدوير الملائمة، لذلك ينبغي أن تكون صياغة الفرضية أو التساؤل البحثي دقيقة ومحددة وكاملة، كما ينبغي أن تكون صياغة الفرضية أو التساؤل البحثي – تمكن الباحث أو غيره من الباحثين من تقييم تصميم الدراسة وإجراءاتها ونتائجها.

(٣) **تحديد نمط التحليل العاملي المناسب**: إن اختيار نمط التحليل العاملي يعتمد علي الهدف من الدراسة العاملية، وعلي مصفوفات الارتباطـــات، وموضـــع تركيـــز التحليل العاملي سواء علي الارتباطات بين الـــصفوف (درجـــات الأفـــراد) أو

الأعمدة (المتغيرات). ويمكن الرجوع إلى الجزء الخاص بأنماط التحليـــل العـــاملي الذي سبقت مناقشته لكي يسترشد به الباحث في تحديد النمط المناسب لدراسته.

٤) **تحديد نموذج التحليل العاملي المناسب:** هناك نموذجان للتحليل العاملي همـــا: نموذج المكونات الرئيسة Principal Factor Component Model، ونمـــوذج التحليل العاملي الطائفي Common Factor. والنمـــوذج الأول لا يـــستند إلى افتراضات فيما يتعلق بالأخطاء النوعية، أو أخطاء الصدفة في البيانات، ويتم تحليل التباين الكلي بعامة، وتكون العامل المستخلصة تركيبا خطيا من المتغيرات، وهنـــا يوضع الواحد الصحيح في الخلايا القطرية لمصفوفة الارتباطات، حيث إنه يعـــد بمثابة الاشتراكيات Communalities. لكن نموذج التحليل العـــاملي الطـــائفي يفترض تقسيم التباين الكلي للمتغير إلى مكونين هما: التباين المشترك، والتبـــاين المتفرد، ويقسم التباين المتفرد إلى تباين نوعي، وتباين الخطأ. لذلك يهـــتم هـــذا النموذج بتحديد العوامل التي تفسر التباين المشترك بين المـــتغيرات. ونظـــراً لأن التباين المشترك في المتغير لا نستطيع تحديده لحين استخلاص العوامل المشتركة أو الطائفية، فإن الاشتراكيات لا تكون محددة، وإنما ينبغي تقديرها. وقد لاحظنا أن التحليل العاملي ألفا Alpha Factor Analysis، والتحليـــل العـــاملي لـــصور المتغيرات Image Factor Analysis يعدان من نماذج التحليل العاملي الطائفي، لذلك ينبغي أن يختار الباحث النموذج الذي يناسب التساؤل البحثي، ويسترشد بخصائص النموذج في اتخاذ قرارات بشأن عدد العوامل.

٥) **تحديد متغيرات الدراسة، وعينة الأفراد أو الكينونات أو الوحدات:**

تعتمد مصفوفة الارتباطات اعتماداً أساسياً على مجموعة المتغيرات التي انتقاهـــا الباحث وعينة الأفراد التي طبق عليها الاختبارات والمقاييس المتعلقة بهذه المتغيرات، أو عينة الكينونات أو الوحدات التي جمع بيانات تتعلق بها في ضوء هذه المتغيرات، ونقصد بالكينونات Entities قضايا اجتماعية أو سياسية،أما الوحـــدات Units فقد تكون وحدات مكانية أو زمنية وغير ذلك.

ولذلك ينبغي أن يحدد الباحث أسلوب المعاينات الذي يستخدمه، وأساليب قياس المتغيرات ومستوي القياس، أي ما إذا كان اسميا أو رتبيا أو فتريا أو نسبيا – كما سبق أن أوضحنا – وكيفية جمع البيانات، وما إذا كانت تعتمد علي تطبيق اختبارات ومقاييس، أو استبيانات أو تقديرات، أو تحليل محتوي وثائق ومطبوعات، وشكل توزيع البيانات التي سيجري تحليلها. لذلك فإن تحديد طبيعة مصفوفة البيانات، وأسلوب المعاينات، ومستوي قياس المتغيرات، وأسلوب جمع البيانات تعد جوانب أساسية متفاعلة ومتكاملة، وليست منفصلة أو متسلسلة، فاتخاذ الباحث قرارا بشأن أي جانب منها يؤثر بالفعل في قراره بشأن الجوانب الأخرى.

وإلى جانب القرارات المتعلقة بمصفوفة البيانات ينبغي أن يتخذ الباحث قرارات تتعلق بإجراءات التصميم. ونقصد بذلك أسلوب التحليل العاملي المستخدم، وطريقة تدوير العوامل، واختبارات الدلالة الإحصائية. إن القرارات التي يتخذها الباحث في هذه الحالات تمكنه من اختيار الطرق والأساليب المناسبة لطبيعة البيانات التي قام بجمعها، ومستوي قياسها، والافتراضات التي يستند إليها الأسلوب أو الطريقة التي سوف يستخدمها في التحليل العاملي. والحقيقة أن العمليات الإجرائية للبيانات والتصميم تعد من العمليات التفاعلية، وتتطلب استبصاراً واعياً بالهدف من الدراسة وتساؤلاتها، والإطار النظري الذي تستند إليه، ومتطلبات نموذج التحليل العاملي المستخدم.

(٦) **تكوين مصفوفة الارتباطات:** بعد جمع البيانات بأساليب منظمة ومتسقة وصادقة يمكن أن يقوم الباحث بتشفير هذه البيانات Data Coding بما يناسب التحليل باستخدام الحاسوب. وهنا ربما يواجه الباحث مشكلة البيانات الناقصة Missing Data بسبب غياب الأفراد، أو عدم انطباق أسئلة أو استفسارات معينة عليهم، أو ترك فقرات دون إجابة سواء في الاستبيانات أو الاختبارات والمقاييس. هنا ربما يلجأ الباحث إلى استخدام أحد الأساليب التالية للتغلب علي هذه المشكلة:

(أ) إذا وجد أن أحد المتغيرات يرتبط ارتباطاً مرتفعاً بمتغير آخر وحصل الفرد علـــي درجة في هذا المتغير الآخر، فإنه يمكن التنبؤ بدرجتـــه الناقـــصة في المـــتغير الأول باستخدام أسلوب الانحدار الخطي البسيط. وإذا كانت هناك درجتان ناقـــصتان، فإنه يمكن التنبؤ بهما بمعلومية درجة متغير مرتبط بكل منهما باستخدام أســـلوب الانحدار المتعدد. غير أن إجراء تحليل الانحدار ربما يستغرق وقتاً وجهداً كـــبيرين، علي الرغم من أنه يُعد أفضل الأساليب للتغلب علي مشكلة البيانات الناقصة.

(ب)يمكن أن يعين الباحث قيمة متوسطة درجات المتغيرات للمتغير الخالي من الدرجة، إذا لم يكن لديه معلومات أخرى.

(ج) استبعاد الفرد الذي تكون درجاته ناقصة في أي من المتغيرات مـــن مـــصفوفة الارتباطات. وعلي الرغم من أن كل قيمة في هذه المصفوفة سوف تعتمد على عدد مختلف من الأفراد، إلا أن كثيراً من الباحثين يلجئون إلى هذا الأسلوب، لكن ينبغي أن يدرك الباحث أن ذلك يؤثر في تفسير الاشتراكيات والجذور الكامنـــة في المـــصفوفة العاملية، فعندما يتم استبعاد بعض الأفراد نجد أن قيمة الاشتراكية ٠,٨٠ مثلاً لا تشير إلى التباين الذي يمكن تفسيره بالنسبة للعدد الكلي للأفراد، وإنما إلى التباين الذي يمكن تفسيره بالنسبة إلى عدد أقل من الأفراد، وهكذا. وعموماً، فإنه ينبغي علي الباحـــث اختيار عدد كبير نسبياً من الأفراد لكي لا تؤثر البيانات الناقصة في نتائج التحليـــل، ويمكن للباحث بعد ذلك تكوين مصفوفة الارتباطات بين المتغيرات لعينة الأفـــراد أو الكينونات أو الوحدات، علماً بأن الكثير من برامج الحاسوب الجاهزة تكون هـــذه المصفوفة باستخدام البيانات الخام مباشرة.

(٧) تحديد أسلوب التحليل العاملي والتوصل إلى مصفوفة التشبعات العاملية:

مع تعدد طرق التحليل العاملي وانتشار البرامج الإحصائية الجاهزة أصبح بالإمكان اختيار الطرق والأساليب العاملية بسهولة،ويشيع استخدام أسلوب المحاور الرئيـــسة Principal Axes Technique في إجراء التحليل العاملي لمصفوفات الارتباطات، غير أن اختيـــار الباحث الأسلوب المناسب يعتمد علي وحدات ومستويات قياس المتغيرات، والأســـاليب

التي استخدمتها الدراسات الأخرى في المجال الذي يهتم به الباحث، حيث إن الأسلوب الشائع الاستخدام في تلك الدراسات ربما يستخدمه الباحث بثقة أكبر، ويمكنه من مقارنة نتائجه بنتائج تلك الدراسات؛ نظراً لأن الأسلوب لم يتغير،كما أن إمكانية تحليل البيانات باستخدام الحاسوب يعتمد علي توافر حزم البرامج الإحصائية وفاعلية إدارتها، ومن الجدير بالذكر أن تحديد طريقة أو أسلوب التحليل العاملي ينبغي أن يسترشد بالقرارات التي اتخذها الباحث في الخطوات السابقة.

(٨)تدوير مصفوفة التشبعات العاملية:

بعد تحديد طريقة أو أسلوب التحليل العاملي والتوصل إلى مصفوفة التشبعات العاملية، يمكن أن يلجأ الباحث إلى تدوير هذه المصفوفة Factor Matrix Rotation. فالمصفوفة العاملية قبل التدوير ربما تكون مصفوفة وحيدة مناظرة لمصفوفة الارتباطات، بينما يوجد عدد لا نهائي من المصفوفات المناظرة بعد التدوير، غير أن الباحث قد يتوصل إلى مصفوفة عاملية تحقق محكات التكوين البسيط Simple Structure التي يتم الحصول عليها باستخدام طريقة تدوير مناسبة.

في كل الأحوال ينبغي أن يتخذ الباحث قرارا بشأن الطريقة التي يستخدمها في التدوير، أي هل يختار طريق بيانية أم طريقة تحليلية. فالطريقة البيانية تعتمد علي التمثيل البصري للنقط والمحاور المرجعية في فضاء متعدد الأبعاد. فالنقط تمثل تشبعات المتغيرات بكل زوج من العوامل، ويتم تدوير محور العامل بصرياً حول نقطة الأصل بحيث يتضح الشكل التكويني للمتغيرات Variable Configuration، ويصبح أكثر قابلية للتعريف في ضوء محكات التكوين البسيط.

غير أنه إذا كان عدد المتغيرات أو العوامل كبيراً تصبح الطريقة البيانية للتدوير غير مناسبة، عندئذ ينبغي أن يستخدم الباحث إحدى الطرق التحليلية، وبخاصة الطريقة Varimax التي اقترحها كايزر Kaiser، والتي يتم إجراؤها باستخدام برامج الأسلوب. وهذه الطريقة تعتمد في تدوير المحاور المرجعية تدويراً متعامداً علي تعظيم أو تصغير دالة رياضية معينة، وهذه الطريقة شائعة الاستخدام. أما إذا لجأ الباحث إلى التدوير المائل لأسباب تتعلق

بهدف الدراسة، فإنه يفضل أن يستخدم التدوير المتعامد أولا ثم يقارن في ضوئه نتائج التدوير المائل. والتدوير المائل الذي يحقق محكات التكوين البسيط يؤدي إلى تجمعات مرتبطة من المتغيرات التي يكون تعريفها أفضل، ويمكن تحديد الارتباط بين التجمعات المختلفة، فإذا كانت هذه التجمعات متعامدة من الوجهة الإمبيريقية، فإن التدوير المائل سوف يؤدي إلى أبعاد متعامدة. ويمكن أن يستخدم الباحث إحدى طرق التدوير المائل، وتؤدي هذه الطرق إلى مصفوفة البنية العاملية، ومصفوفة الارتباطات بين المتغيرات والعوامل.

(٩) تسمية العوامل وتفسيرها:

يقصد بتسمية العوامل اقتراح عناوين موجزة للعوامل المستخلصة سواء قبل التدوير أو بعده. وتعكس هذه العناوين تقييم الباحث للنتائج، وتشبعات تجمعات المتغيرات بالعوامل، ومحتوي أو مضمون هذه المتغيرات.

(١٠) التوصل إلى تقديرات درجات العوامل:

قد تتطلب الدراسة التوصل إلى تقدير درجات العوامل، حيث يمكن استخدام هذه التقديرات في تصميم بحوث أخرى، مثل البحوث التي تعتمد علي الانحدار المتعدد أو تحليل التباين، أو دمجها مع متغيرات أخرى وإجراء تحليل عاملي جديد، أو المقارنة بين العوامل التي يتم استخلاصها من دراسات مختلفة في مجال معين.

عاشراً:صياغة نتائج الدراسة العاملية

بعد الانتهاء من إجراء التحليل العاملي يقوم الباحث بكتابة تقرير دراسته لكي يطلع عليه الباحثون الآخرون والمهتمون بنتائج الدراسة. لذلك ينبغي أن يتضمن التقرير معلومات كافية تسمح لهؤلاء بالتحقق من نتائجها. إذ ليس كافياً أن يعرض الباحث النتيجة النهائية فقط، بل ينبغي أن يقدم بياناته الأصلية لكي يتسنى متابعة التحليل من بدايته. وإذا لم يكن هذا ممكناً، فإن الباحث يمكنه تقديم مصفوفة الارتباطات التي استخدمت في التحليل العاملي، وبذلك يمكن للباحثين الآخرين إعادة التحليل إذا رغبوا في ذلك، أو استخدام أساليب إحصائية أخرى أكثر ملاءمة لاستخلاص مزيد من المعلومات من هذه المصفوفة.

وعلي الرغم من أن كثيراً من الدوريات العلمية لا تسمح بنشر مصفوفات الارتباطات، إلا أنه يمكن تخزينها في الحاسوب لمن يطلبها. وكذلك ينبغي تقديم المصفوفات العاملية قبل التدوير وبعده وغيرها من المصفوفات، وبخاصة إذا كانت الدراسة لنيل درجة الماجستير أو الدكتوراه أو للترقية العلمية. غير أن بعض الذين يطلعون علي الدراسات العاملية ربما يودون فقط تقييم مدي كفاية الأساليب التي استخدمها الباحث، والإجراءات التي اتبعها في التحليل العاملي لبيانات دراسته، لذلك ينبغي أن يزود الباحث هؤلاء بمعلومات كافية تمكنهم من هذا التقييم،ويمكن أن يسترشد الباحث بالعناصر الأساسية التالية في كتابة التقرير لكي يكون التواصل العلمي فاعلا:

(١) **متغيرات الدراسة:** فالمتغيرات التي اختارها الباحث ينبغي أن يقوم بوصفها بالتفصيل لكي يوضح للقارئ ماذا تقيسه هذه المتغيرات والأساس الذي استند إليه في اختيارها، وكذلك يبين أدوات القياس التي استخدمها من حيث المحتوى والثبات والصدق، وكيفية تصحيحها أو تقدير درجاتها، وظروف تطبيقها، وتوزيع درجاتها، والقيم الإحصائية الوصفية لهذا التوزيع، مثل النزعة المركزية، والتشتت، والالتواء، والتفرطح.

(٢) **عينة الدراسة:**وهنا ينبغي أن يوضح الباحث المجتمع المستهدف الذي استمدت منه العينة، وخصائص هذا المجتمع الديموجرافية، وأسلوب المعاينات المستخدم، وعدد أفراد العينة بالنسبة لعدد المتغيرات التي استخدمها(نظرا لأن حجم العينة يؤثر في قيم معامل الارتباط وبالتالي يؤثر في قيم التشبعات العاملية).

(٣) **معامل الارتباط المستخدم:** إن معاملات الارتباط والاقتران تعتمد على مستوى قياس كل من المتغيرين، لذلك ينبغي أن يوضح الباحث نوع معامل الارتباط المستخدم وطريقة حسابه،فعلى الرغم من أن معامل ارتباط بيرسون هو أكثر هذه المعاملات استخداما في التحليل العاملي لأنه يعتمد علي متغيرات من المستوى الفتري والنسبي بما يتفق ومتطلبات التحليل العاملي، إلا أن بعض الباحثين ربما

يستخدم معامل الارتباط الثنائي المتسلسل، كذلك ينبغي أن يوضح ما إذا كان أجرى تحويلاً معيناً على درجات المتغيرات قبل حساب معامل الارتباط.

(٤) **استخلاص العوامل:** يتطلب أن يوضح الباحث نموذج التحليل العاملي، والأسلوب الذي استخدمه في استخلاص العوامل من مصفوفة الارتباطات، والمحك الذي استخدمه في تحديد عدد العوامل، وكذلك يوضح مصفوفة التشبعات العاملية المستخلصة.

(٥) **الاشتراكيات:** فالخلايا القطرية في مصفوفة الارتباطات ينبغي أن توضع فيها قيم الاشتراكيات، ومن المعروف أن هذه القيم يمكن أن تكون محددة بالواحد الصحيح كما في أسلوب المكونات الرئيسة، أو قيم تقديرية كما في أساليب التحليل العاملي الأخرى، لذلك ينبغي أن يوضح الباحث كيف تم تقدير هذه الاشتراكيات، وما إذا كان قد استخدم عمليات التكرار المتتالي Iteration للتوصل إلى قيمها التقديرية.

(٦) **تدوير مصفوفة التشبعات العاملية:** يفترض أن يوضح الباحث ما إذا كان التدوير متعامداً أو مائلاً، جودة نتائج التدوير، وتبرير هذه المحكات، وكذلك يوضح المصفوفة العاملية بعد التدوير.

حادي عشر:استخدام برامج الحاسوب في إجراء التحليل العاملي:

لإجراء التحليل العاملي يمكن استخدام أحد برامج الحاسوب المتوافرة ضمن الحزم الإحصائية SPSS، أو BMD، أو SAS. ونظراً لأن كثيراً من الباحثين يعتمدون علي الحزمة SPSS، فسوف نلقي الضوء علي البرنامج الذي يجري التحليل العاملي في هذه الحزمة. فالبرنامج FACTOR هو أحد البرامج الفرعية في هذه الحزمة، ويجري التحليل العاملي بمعظم الأساليب المعاصرة التي أوضحناها.

مدخلات البرنامج:

(١) البيانات الخام، أو مصفوفة الارتباطات، أو مصفوفة التغاير Covariance Matrix.

(٢) يحدد الباحث عدد العوامل التي يود استخلاصها من البيانات، أو محك تحديد هذا العدد.

(٣) يحدد ما إذا كان يود وضع القيم التقديرية للاشـــتراكيات في الخلايـــا القطريـــة لمصفوفة الارتباطات ونوع التقديرات المطلوبة.

(٤) يحدد ما إذا كان يود تدوير المصفوفة العاملية تدويراً متعامداً أم مائلاً، وطريقـــة التدوير التي يختارها.

(٥) يمكن أن يحدد عدد مرات التكرار المتتـــالي للعمليـــات Iteration للتوصـــل إلى التقارب Convergence.

مخرجات البرنامج:

(١) المتوسطات والانحرافات المعيارية لجميع المتغيرات المراد تحليلها

(٢) مصفوفة الارتباطات

(٣) معكوس مصفوفة الارتباطات ومحدداتها

(٤) مصفوفة التشبعات العاملية قبل التدوير

(٥) مصفوفة التشبعات العاملية بعد تدويرها بالطريقة Varimax

(٦) مصفوفة درجات العوامل متضمنة أوزان الانحدار التي تستخدم في بناء مـــوازين العوامل Factor Scales.

(٧) تمثيل بياني للعوامل بعد تدويرها متضمناً أشكالاً لتجمعات المتغيرات لكـــل زوج من العوامل.

ثاني عشر:أبرز الأخطاء الشائعة في استخدام التحليل العاملي:

يستخدم كثير من الباحثين أساليب العاملي في دراستهم سواء للتحقق من الصدق العاملي Factorial Validity لأدوات الدراسة، أو للكشف عن أقل عدد من العوامل التي تفسر مجموعة من المتغيرات المتعلقة بظاهرة معينة يهتمون بدراستها، أو التحقـــق مـــن صـــحة

فرضيات معينة. غير أنه ربما يُساء استخدام التحليل العاملي، وذلك بسبب قلة الخبرة وقلة فهم هذا التحليل وما يستند إليه من نماذج إحصائية ورياضية متقدمة، وبخاصة أن كثيراً من الباحثين لم ينل القدر الكافي من دراسة الرياضيات العالية مثل جبر المصفوفات، وهندسة المتجهات، وعلم التفاضل، لذلك نجد الكثير من الأخطاء الشائعة في استخدام التحليل العاملي، يتمثل أهمها فيما يلي:

(١) إجراء التحليل العاملي دون الاستناد إلى تصميم علمي دقيق يأخذ بعين الاعتبار الخطوات العشر التي أوضحناها فيما سبق، وإنما يقوم الباحث مباشرة بجمع البيانات، ويستخدم برامج الحاسوب الجاهزة في إجراء التحليل دون دراية كافية بهذه البرامج ومتطلباتها ونوعية مخرجاتها وما تحتوي عليه، مما يجعل التحليل العاملي للبيانات عملية آلية تفقد البيانات أهم ما تنطوي عليه من معلومات.

(٢) استخدام عدد كبير من المتغيرات التجريبية في التحليل العاملي ليس بسبب أهميتها وإنما لتوافرها لدى الباحث؛ مما يؤدي إلى تعقيد إجراءات تدوير العوامل وتفسيرها،فعدد المتغيرات ينبغي أن يزيد عدة مرات عن عدد العوامل، فمثلاً ينبغي أن يكن هناك خمسة متغيرات جيدة ومناسبة علي الأقل مبدئيا لكل عامل متوقع استخلاصه.

(٣) عدم التحقق من الافتراضات التي يتطلبها التحليل العاملي في البيانات والمتعلقة بمستوى قياس المتغيرات وشكل توزيعاتها، فبعض الباحثين يستخدم متغيرات توزيعاتها ملتوية التواء شديداً أو مبتورة، أو متعددة المنوال، أو مقسمة ثنائياً تقسيماً متطرفاً أو العلاقات فيما بينها ملتوية.

(٤) استخدام بيانات تتعلق بمتغيرات متداخلة (غير مستقلة) من الوجهة التجريبية، كأن يكون أحد المتغيرات مركب من المتغيرات الأخرى تركيبا خطيا مثل درجات الذكاء، ودرجات الاستعداد اللفظي، والدرجات الكلية، أو أن تكون الاستجابة على بند معين في استبيان تعتمد علي الاستجابة على بند آخر،

وكذلك استخدام متغيرات مركبة ولا تتميز بالبساطة، أي تقيس عامل مشترك واحد بدرجة ما.

(٥) عدم الاهتمام بعدد المتغيرات المشبعة بالعوامل، إذ ينبغي أن لا يقل عدد المتغيرات المشبعة بكل عامل عن ثلاثة متغيرات. فإذا اعتني الباحث بتصميم دراسته، فإنه يمكنه التنبؤ بعدد العوامل المستخلصة، وعما إذا كانت ستتوزع توزيعاً متعادلاً علي العوامل.

(٦) استخدام متغيرات متشابهة في التحليل مما يؤدي إلى استخلاص عوامل من المستوى الأدنى في التنظيم الهرمي للعوامل (فلا يجوز مثلا استخدام صورتين متكافئتين من اختبار أو مقياس معين).

(٧) عدم تصميم خطة انتقاء عينة الأفراد التي ستطبق عليها الاختبارات والمقاييس، فأحياناً تكون العينة صغيرة الحجم، أو غير ممثلة للمجتمع المستهدف، أو متحيزة، أو تكونت نتيجة ضم مجموعتين متمايزتين مختلفتين في البنية العاملية للمتغيرات. وإذا كانت عينة الأفراد غير متجانسة في العمر أو النوع أو مستوى التعليم أو المستوى الاجتماعي الاقتصادي، فإن العوامل المستخلصة تتأثر بذلك

إن اختيار عينة غير متجانسة بالنسبة لهذه المتغيرات يعتمد علي المجتمع المراد تعميم النتائج عليه، فإذا كان تفسير العوامل يتعلق بالفروق الفردية في مستويات عمرية معينة، فإنه ينبغي أن تكون العينة متجانسة نسبياً في العمر. أما إذا كانت العوامل تتعلق بالنزاعات النمائية، فإن العينة ينبغي أن تكون ممثلة لمدى عمري متسع. وإذا استخدمت عينة مشتركة من البنين والبنات، يُفضل تحويل درجات كل منهما إلى درجات معيارية قبل إيجاد قيم معاملات الارتباط، وفي هذه الحالة ينبغي إدخال (النوع) كمتغير في عمليات التحليل، وعزل أثره، ثم إجراء التحليل العاملي علي مصفوفة البواقي.

(٨) قلة عدد العوامل المستخلصة نتيجة عدم وجود عدد كبير من النقط في الفضاء متعدد الأبعاد، مما يؤدي إلى صعوبات في تدوير العوامل لكي تحقق محكات

التكوين البسيط. لذلك ينبغي أن يصمم الباحث دراسته بحيث يكون عدد المتغيرات كافياً لاستخلاص خمسة أو ستة عوامل متعامدة نسبياً علي الأقل.

(٩) استخدام معاملات ارتباط غير مناسبة مثل معامل فاي، أو معامل الارتباط الرباعي دون التحقق من عدم مخالفة هذا المعامل للافتراضات التي يستند إليها في البيانات.

(١٠) استخدام قيم اشتراكيات غير مناسبة في الخلايا القطرية لمصفوفة الارتباطات، كأن يضع الواحد الصحيح في هذه الخلايا عند استخدام التحليل العاملي الطائفي Group Factor Analysis. فالواحد يصلح إذا استخدم الباحث أسلوب المكونات الرئيسة Principal Components.

(١١) استخدام طرق غير مناسبة في تدوير المصفوفة العاملية، أو استخدام التدوير المتعامد عندما يكون التدوير المائل أكثر ملاءمة لإعطاء نتائج أفضل، أو عدم الاستناد إلى محكات في عملية التدوير.

(١٢) تفسير العامل الأول الذي يتم استخلاصه علي أنه عامل عام، وإعطاء تسمية للعوامل دون فحص طبيعة هذه العوامل، ومحتوي المتغيرات المتشبعة بها، وغير ذلك من أسس تسمية العوامل وتفسيرها، وهذا يتطلب التحقق بأساليب أخرى من صدق التكوين الفرضي لهذه العوامل.

(١٣) استخدام التشبعات العاملية في المقارنة بين العوامل المستخلصة من تحليلين مختلفين. فإذا أراد الباحث تحديد التماثل بين العوامل من دراستين مختلفتين لا بد أن يكون واعياً بحقيقة أن العوامل تُعد تركيبات خطية من المتغيرات، وهذه التركيبات يتم تحديدها بالفعل في حالة المكونات الرئيسة، أو تقديرها كما في التحليل العاملي الطائفي،وهذه التركيبات الخطية ينبغي التمييز بينها وبين التشبعات العاملية التي تدل علي الارتباط بين المتغيرات والعوامل، فالتشبعات بهذا المنظور ليست عوامل يتم مقارنتها.

والطريقة الصحيحة للمقارنة بين العوامل في الدراسات المختلفة هو أن نقارن بين القيم التقديرية لدرجات العوامل Factor Scores Estimates في الدراستين مع مراعاة أن تكون الدراستان قد استخدمتا عينة الأفراد نفسها حتى لو اختلفت المتغيرات،فعندئذ يمكن إيجاد الارتباط بين درجات العوامل في كل من الدراستين، وبذلك يمكن المقارنة بين العوامل في ضوء قيم معامل الارتباط الناتجة. أما إذا اختلفت عينات الأفراد دون أن تختلف المتغيرات فيمكن المقارنة بين العوامل بتعيين أوزان معينة للمتغيرات في الدراستين المختلفتين، وإيجاد القيم التقديرية لدرجات العوامل، ثم إيجاد قيم معامل الارتباط بين مجموعتي الدرجات في الدراستين.

وإذا اختلفت الدراستان في عينة الأفراد والمتغيرات فإنه في هذه الحالة يكون من الخطأ أن تتم المقارنة بين العوامل المستخلصة في كل منهما(يمكن الحصول على درجات العوامل باستخدام معادلات الانحدار المتعدد غيرها من الأساليب)

الفصل السادس
إحصاءات العينة

تقديم:

من المعروف أن العينة جانب جوهري في البحث العلمي بكافة فروعه وتطبيقاتـــه، وقـــد اهتمت أدبيات علم الإحصاء بالعينات من زوايا متعددة، بل إن العينات تخصص عميـــق ضمن هذا العلم . الفصل الحالي يتضمن أهم الجوانب الإحصائية في العينات، حيث يقدم تعريفاً موجزاً بالعينات من حيث مفهومها وخصائصها بما في ذلك حجم العينة والعوامل المؤثرة فيه والطرق الإحصائية لتقديره، وكيفية تحديده مع توضيح كيفية استخدام الجداول الإحصائية التي تساعد الباحثين في هذا الشأن، كما يتضمن الفصل تعريفـــاً بـــإجراءات المعاينة العشوائية، وكذلك توضيح أخطاء المعاينة لما لها من دلالة فائقة الأهمية ونحن بصدد التقدير والاستدلال على خصائص المجتمع بناءً على دراسة العينة

المبحث الأول
حجم العينة وطرق تقديره

أولاً: مفهوم العينة وخصائصه:

العينة (Sample) هي جزء من المجتمع، بحيث تتوافر في هذا الجزء نفس خصائص المجتمـــع. والحكمة من إجراء الدراسة على العينة- تتمثل في أنه في كثير من الأحيان يستحيل إجراء الدراسة على المجتمع، وبالتالي يتم اختيار عينة محدودة ودراستها بهدف التوصل إلى نتـــائج يمكن تعميمها على المجتمع. ويصبح ذلك ممكناً، إذا كانت خصائص العينة تمثل خصائص المجتمع من حيث أكبر عدد ممكن من المتغيرات، خاصة المتغيرات التي يحتمل أن تـــؤثر في الظاهرة محل البحث. افترض أن مجتمع البحث عدة آلاف من التلاميذ وليكن عشرين ألفاً، في هذه الحالة قد يكون من غير الممكن إجراء الدراسة على هذا العدد، وهنا يتم اختيـــار عينة من التلاميذ بحيث تتوافر فيها خصائص المجتمع، فإذا كان مجتمـــع البحـــث يتـــوزع حسب متغير الجنس – بين الذكور بنسبة ٦٠% والإناث بنسبة ٤٠%، فـــإن الباحـــث يمكنه أن يجعل العينة تعكس هذا التوزيع وتكون في الوقت نفسه ممثلة للمجتمع باستخدام طرق المعاينة العشوائية، فإذا كان حجم العينة هو (١٠٠) مفردة مثلاً، فهذا يعني أن العينة سوف تتضمن (٦٠) من الذكور، مقابل (٤٠) من الإناث ... وهكذا في بقية المتغيرات التي قد تتمثل في الصف الدراســـي، ومحل الإقامة، ومستوى دخل الأسرة ... الخ.

ويفترض أن تكون العينة ممثلة للمجتمع الذي سحبت منه، ويتم ذلك من خلال الاختيار العشوائي حتى يمكن تعميم نتائج البحث، لكن هناك ظروفاً معينـــة، بـــل ومنـــاهـــــج بحـــثية (كمنهج دراسة الحالة)، وكذلك الدراسات الاستكشافية يمكن فيها إجراء الدراسة

على عينات غير احتمالية، أي غير عشوائية، وبالتالي يتم التعامل مع النتــائج في حــدود معينة، أو لغرض معين. من جهة أخرى، فإن حجم العينة، يختلف من بحث إلى آخر، كما تتعدد أساليب اختيار العينات، الأمر الذي يتضح في النقاط التالية.

ثانياً: حجم العينة:

يقصد بحجم العينة عدد المفردات التي ستجرى عليها الدراسة، وليس هناك حجم ثابــت يصلح لجميع الدراسات، فبعض البحوث تجرى على بضعة أفراد، أو عشرات أو مئات أو ألوف الأفراد، وقد لوحظ في كثير من المؤلفات التي تناولت مسألة حجم العينة — العديد من المغالطات فيما يتعلق بالحد الأدنى والحد الأقصى لحجم العينة؛ كالقول مثلاً بأن الحد الأدنى للعينة في الدراسات الوصفية يجب أن يكون (١٠%) من المجتمع الأصلي، فهذا غير صحيح بالمرة، ماذا نقول مثلاً إذا كان حجم المجتمع الأصلي عشرة ملايين نسمة مــثلاً، فهل يكون حجم العينة مليون شخص؟ إنه من الصعب، إن لم يكن من المستحيل إجــراء دراسة على عينة بهذا الحجم الضخم. إن حجم العينة يتم تحديده على ضوء أكثــر مــن اعتبار أهمها: طبيعة المجتمع الأصلي خاصة من حيث العدد ومن حيث تجانس المفــردات، فإذا كان المجتمع الأصلي — على سبيل المثال — هو التلاميذ مكفوفــو البــصر، والبــالغ عددهم مائة تلميذ وتلميذة، فإن الدراسة قد تشملهم جميعاً وقد تشمل خمسين أو ثلاثين منهم، أما إذا كان حجم المجتمع الأصلي مليون فرد، فإن حجم العينة يمكن أن يكون بضع مئات من المفردات، أو ألف مفردة. كما أن مدى التجانس بين أفراد المجتمع من العوامل المؤثرة في تحديد حجم العينة، فكلما كان المجتمع أقل تجانساً (أي أكثر تبايناً) يصبح مــن الضروري اختيار عينة كبيرة نسبياً حتى نضمن أن تشتمل العينة على عدد كــاف مــن المفردات بما يضمن تمثيل المفردات المتباينة، أما إذا كان التباين صغيراً (بمعــنى أن يكــون المجتمع أكثر تجانساً) فإنه من الممكن تصغير حجم العينة. (من أمثلة المجتمعــات الأكثــر تجانساً (الأقل تبايناً) تلاميذ صف دراسي معين في نمط معين من المدارس ولتكن مدارس اللغات، ومن أمثلة المجتمعات الأقل تجانساً (الأكثر تبايناً) مجتمع المدن بفئاته المختلفة

العامل الآخر المؤثر في تحديد حجم العينة هو أدوات جمــع البيانات،فكلمــا تعــددت الأدوات (والتي ستطبق على كل فرد) فإن ذلك يستدعي تصغير حجم العينة،أما إذا كانت الدراسة تستخدم أداة واحدة، أو أدوات محدودة فيمكن زيادة حجم العينــة، وفي كــل الأحوال فإن حجم العينة يجب أن يتفق ومتطلبات الحصول على بيانات كافية ونتــائج موثوق فيها. ويتأثر حجم العينة كذلك بعامل التكلفة ،فبعض البحوث تتطلب تكلفة مالية وتكنولوجية، وسفر وانتقالات ... الخ، بما يفرض أن يكون حجم العينة في ضوء المــوارد المتاحة، فكلما كانت التكلفة مرتفعة بحيث يتعذر إجراء الدراسة على عينة من ألف مفردة

مثلاً، يصبح تخفيض حجم العينة أمراً ضرورياً بحيث يكون في حدود الموارد المتاحة. كما أن عامل الزمن أو الوقت يؤثر في حجم العينة، فهناك بحوث يتعين إنجازها خلال فترة زمنية معينة، فإذا كان حجم العينة كبيراً فإن الدراسة لن تنجز في الوقت المحدد، وقد تستغرق الدراسة زمناً طويلاً لإنجازها وحينئذ يمكن أن تكون الظاهرة المدروسة قد تغيرت بصفة كلية أو جزئية فتقل القيمة العلمية للدراسة وقد يتم إلغاء الدراسة إذا لم يتم إنجازها خلال فترة زمنية محددة. كما أن العوامل المتعلقة بطبيعة البحث والهدف منه تؤثر بالتأكيد في حجم العينة، فقد يكون البحث ذا صفة استطلاعية، كأن يستهدف استطلاع آراء مشاهدي التليفزيون حول مستوى البرامج والمسلسلات، وهنا يمكن أن يكون حجم العينة بضع مئات أو حتى بضع ألوف، أما عندما يكون البحث تجريبياً بهدف معرفة أثر طريقة جديدة في تقديم برامج التليفزيون،فإن حجم المجموعة الواحدة يتحدد حسب رؤية الباحث وتقييمه، فقد يختار الباحث مجموعة من ثلاثين فرداً كمجموعة تجريبية، ومجموعة أخرى مماثلة لها في العدد كمجموعة ضابطة. ويلاحظ هنا أن طبيعة الدراسة التجريبية تفرض أن يكون حجم العينة صغيراً بما يتناسب مع المكان ومستلزمات التجربة. ويتأثر حجم العينة كذلك بتوقعات عدم الاستجابة، فكلما زادت التوقعات بأن يرفض بعض المبحوثين المشاركة في البحث، أو الإجابة على أسئلة معينة، يصبح من الضروري أخذ ذلك بالاعتبار وزيادة حجم العينة، فقد يقرر الباحث أن يجري دراسته على عينة قوامها (١٠٠٠) مفردة، لكنه يرى أن من المحتمل أن يرفض بعض الأفراد المشاركة في البحث، أو يرفضون الإجابة على أسئلة معينة، ومن هنا يجري دراسته على ١٠٥٠ مفردة ، أي بزيادة قدرها خمسون مفردة عن العدد المقرر، وإذا أردنا إجراء بحث على ٢٠٠٠ حالة مثلاً وكان تقديرنا أن ٢٠% من الحالات سوف لا تستجيب، فإنه من المناسب أن نبدأ بعينة حجمها ٢٥٠٠ حالة (علما بأن ذلك لا يقلل من أخطاء التحيز).

كما أن حجم العينة يتأثر بمستوى الدقة المطلوب، فمن المعروف أنه كلما صغر حجم العينة زاد تأثر النتائج بعامل الصدفة، وكلما زاد تأثير عامل الصدفة انخفضت الثقة في النتائج، ويتأثر ذلك بمستوى الثقة، ودرجة الدقة المطلوبة، وحدود الخطأ المسموح به. فكلما أراد الباحث مستوى ثقة مرتفعاً، ودرجة دقة عالية،وهامش خطأ ضئيلاً- يتطلب الأمر اختيار عينة أكبر، وذلك مقارنة بما إذا اكتفى الباحث بمستوى ثقة أقل،ودرجة دقة منخفضة مع السماح بهامش خطأ أكبر.

في ضوء هذه الاعتبارات يتم تحديد الحجم المناسب للعينة، ويمكن أن يقتدي الباحث بأحجام العينات التي استخدمت في دراسات مشابهة لدراسته، كما يمكنه الاستعانة بآراء المتخصصين. وهناك نقطة جوهرية يجب أن يعيها الباحثون وهم بصدد تحديد حجم العينة، فالعينة ذات الحجم الأكبر لا تعنى بالضرورة أن تكون نتائجها موثوقا فيها بدرجة أكبر

مقارنة بنتائج مستمدة من عينة صغيرة الحجم، ذلك أن الثقة في نتائج العينة تتوقف على طريقة سحب العينة، أي أن يتم اختيار العينة بطريقة صحيحة بحيث تكون ممثلة للمجتمع وتعكس ما فيه من تباين.

ثالثاً: بعض الطرق الإحصائية لتقدير حجم العينة:

هناك بعض المعادلات الرياضية لتقدير حجم العينة المطلوب سحبها من المجتمع غير المحددة، والتي يصل تعدادها إلى الألوف أو الملايين، ومن بين تلك المعادلات تلك التي ترتكز على تقدير الخطأ المعياري بنسبة ٠,٥٠ وتتمثل تلك المعادلة في:

$$د ق = د ع \sqrt{\frac{ح \, (١ - ح)}{ن}}$$

حيث إن:

د ق : درجة الدقة، ويحددها الباحث بناء على قناعاته، كأن يقول مثلاً بأنه يرتضي أن تكون درجة الدقة ± ٥% بمعنى أنه إذا كشفت الدراسة عن أن نسبة مشاهدة التليفزيون هي ٧٥% فإن هذه النسبة سنتعامل معها على أنه تقل ٥% أو تزيد ٥%، أي أن النسبة ستكون في حدود من ٧٠% إلى ٨٠%

د ع : الدرجة المعيارية المقابلة لمستوى الثقة (إذا استخدمنا مستوى الثقة ٩٥% فإن الدرجة المعيارية المقابلة له هي ١,٩٦ أما إذا استخدمنا مستوى الثقة ٩٩ % فإن الدرجة المعيارية المقابلة له هي ٢,٥٨)

ن : حجم العينة المطلوبة

ح : قيمة الخطأ المعياري للنسبة، وهذه القيمة تساوي ٠,٥ وبالتالي فإن (١- ح) تساوي ٠,٥ أي أن القيمة الموجودة تحت علامة الجذر التربيعي في المعادلة السابقة يتم التعويض عنها بالقيمة (٠,٥× ٠,٥) وذلك باعتبار أن النسبة (٠,٥ × ٠,٥) تعطى أكبر النتائج دقة قياساً على أي نسبتين أخريين (مثال ذلك:(٠,١ × ٠,٩) أو (٠,٢ × ٠,٨) أو ٠(٠,٣× ٠,٧)، أو ٠,٤×٠,٦) ...الخ.

فإذا حددنا مستوى الدقة بأنه ± ٥% وحددنا مستوى الثقة بأنه ٠,٩٥ وبالتعويض في المعادلة المذكورة يمكننا تحديد حجم العينة المطلوبة كالآتي:

$$٥\% = ١,٩٦ \sqrt{\frac{ح(١-ح)}{ن}}$$

$$٥\% = ١,٩٦ \sqrt{\frac{٠,٥ \times ٠,٥}{ن}}$$

بالتخلص من الجذر التربيعي، وبالتعويض فإن:

$$٠,٠٠٢٥ = \frac{٣,٨٤١٦ \times ٠,٥ \times ٠,٥}{ن}$$

٠,٠٠٢٥ ن = ٣,٨٤١٦ × ٠,٥ × ٠,٥

٠,٠٠٢٥ ن = ٠,٩٦٠٤

ن = ٠,٩٦٠٤ ÷ ٠,٠٠٢٥= ٣٨٤,١٦

أي أن حجم العينة المطلوبة هو ٣٨٤ مفردة، وذلك بمستوى ثقة ٠,٩٥ ودرجة دقة ٠,٠٥

أما إذا أخذنا بمستوى الثقة (٠,٩٩) ، ودرجة الدقة ٠,٠٥ فإن حجم العينة المطلوب يتم الحصول عليه بتطبيق المعادلة نفسها مع استخدام الدرجة المعيارية التي تقابل مستوى الثقة٠,٩٩ وهذه الدرجة تعادل ٢,٥٨ أي أن:

$$٥\% = ٢,٥٨ \sqrt{\frac{ح(١-ح)}{ن}}$$

$$٥\ \% = ٢,٥٨ \sqrt{\frac{٠,٥ \times ٠,٥}{ن}}$$

$$٠,٠٠٢٥ = \frac{٠,٥ \times ٠,٥ \times ٦,٦٥٦٤}{ن}$$

$$٠,٠٠٢٥\ ن = ٠,٥ \times ٠,٥ \times ٦,٦٥٦٤$$

$$٠,٠٠٢٥\ ن = ١,٦٦٤١$$

$$ن = ١,٦٦٤١ \div ٠,٠٠٢٥ = ٦٦٥,٦٤$$

أي أن حجم العينة المطلوبة يكون ٦٦٦ مفردة، وذلك عند مستوى ثقة ٠,٩٩ ودرجة دقة ٠,٠٥.

وكثيراً ما يتم إجراء البحوث الإعلامية على عينات بحجم معين بحيث يمكن تقدير متوسط وجود ظاهرة معينة في المجتمع بناء على النتائج المستمدة من العينة، فلنفترض مثلاً أننا نريد تحديد حجم العينة اللازم لتقدير متوسط عدد ساعات استخدام الحاسوب يومياً للقائمين بالاتصال في المؤسسات الإعلامية بالدولة.هنا يتم تحديد مستوى الثقة وليكن ٩٥% ، كما يتطلب الأمر أن نحدد قيمة الخطأ المسموح به، وليكن هنا ± ٠,١٥ من الساعة، ومن الضروري أيضاً تقدير الانحراف المعياري لمتوسط عدد ساعات استخدام الحاسوب يومياً للقائمين بالاتصال في المؤسسات الإعلامية بالدولة، وعادة ما يتم ذلك إما من خلال بحوث سابقة تتضمن قيمة الانحراف المعياري، أو من خلال إجراء دراسة استطلاعية محدودة. لنفرض أن الانحراف المعياري تم تقديره (وفق أسس علمية) بأنه ساعة ونصف (١,٥ ساعة) ، هنا نطبق المعادلة الآتية:

$$٠,١٥ = ١,٩٦ \times \frac{ع}{\sqrt{ن}}$$

حيث إن:

ع : هي الانحراف المعياري للمجتمع

أما $\frac{ع}{\sqrt{ن}}$ فهي الخطأ المعياري للمتوسط الحسابي

وبالتعويض في هذه المعادلة يكون:

$$٠,١٥ = ١,٩٦ \times \frac{١,٥}{\sqrt{ن}}$$

أي أن

$$٠,٠٢٢٥ \text{ ن} = (١,٩٦ \times ١,٥)^{٢}$$

$$٠,٠٢٢٥ \text{ ن} = (٢,٩٤)^{٢} = ٨,٦٤٣٦$$

$$\text{ن} = ٨,٦٤٣٦ \div ٠,٠٢٢٥ = ٣٨٤,١٦$$

أي أن حجم العينة المطلوبة هو ٣٨٤ مفردة

وهناك صيغة أخرى مبسطة للمعادلة السابقة، وهذه الصيغة المبسطة هي:

$$n = \left\{ \frac{Z\,(\sigma S)}{E} \right\}^{2}$$

حيث:

n = حجم العينة

σS = القيمة التقديرية للانحراف المعياري في المجتمع

E = قيمة الخطأ المسموح به

Z = الدرجة المعيارية لمستوى الثقة، وهذه الدرجة تكون ١,٩٦ إذا استخدمنا مستوى الثقة ٠,٩٥ بينما تكون هذه الدرجة ٢,٥٨ إذا استخدمنا مستوى الثقة ٠,٩٩

وباستخدام المعادلة المذكورة فإننا نصل إلى النتيجة نفسها بشأن حجم العينة المطلوبة لتقدير متوسط عدد ساعات استخدام الحاسوب يومياً للقائمين بالاتصال في المؤسسات الإعلامية بالدولة،وكما سبقت الإشارة نفرض أن القيمة التقديرية للانحراف المعياري في المجتمع(σS) هي ١,٥ ساعة، كما أن الخطأ المسموح به (E) يساوي ٠,١٥ ساعة، أما مستوى الثقة فهو ٠,٩٥ والذي يقابله الدرجة المعيارية (Z) وقيمتها ١,٩٦ بهذه المعطيات وباستخدام صيغة المعادلة المبسطة المشار إليها فإن:

$$\text{حجم العينة} = \left(\frac{1,96\ (1,5)}{0,15} \right)^{2} = (19,6)^{2} = 384,16$$

ومن الواضح أنها القيمة نفسها التي تم التوصل إليها باستخدام المعادلة:

$$0,15 = 1,96 \times \frac{\text{ع}}{\sqrt{\text{ن}}}$$

وتغطي المعادلة المبسطة- السابق ذكرها- مدى واسعاً من التطبيقات في التوصل إلى حجم العينة، فإذا أردنا – على سبيل المثال- تحديد حجم العينة المطلوبة لتقدير متوسط قيمة ما تنفقه الأسرة في شراء الصحف اليومية في السنة، فإننا نحدد مستوى الثقة المطلوب، وليكن ٠,٩٥ (أي أننا إذا كررنا التجربة ١٠٠ مرة ، فسوف نحصل على النتيجة نفسها في ٩٥ مرة)، كما نحدد الخطأ المسموح به، وليكن عشرة جنيهات زيادة أو نقصاً ، بمعنى أنه إذا كان متوسط الإنفاق كما ستكشفه العينة يساوي ٣٤٠ جنيها،

فإننا سنعتبره يتراوح مابين ٣٣٠ إلى ٣٥٠ جنيها . أما العامل الأخير والأساسي الذي لا بد من تقديره فهو الانحراف المعياري لمتوسط الإنفاق السنوي على شراء الصحف اليومية في المجتمع (ويتم تقدير الانحراف المعياري في المجتمع بناء على دراسات سابقة أو بناء على عينة محدودة يتم اختيارها بعناية)، لنفترض أنه تم تقدير قيمة الانحـــراف المعيـــاري بأنهـــا (١٥٠) ، بموجب ذلك يكون لدينا : درجة الثقة المطلوبة، الخطأ المسموح بـــه، تقـــدير للانحراف المعياري في المجتمع ، وهنا يمكننا تطبيق المعادلة السابقة كالآتي:

$$\text{حجم العينة} = \left(\frac{١,٩٦\ (١٥٠)}{١٠}\right)^{٢} = ٨٦٤,٣٦$$

أي أن حجم العينة المطلوبة يكون ٨٦٤ مفردة، وذلك بقيمة خطأ مسموح به قـــدرها ١٠ ومستوى ثقة ٩٥% (١,٩٦ وحدة انحراف معياري).

أما إذا استخدمنا مستوى ثقة أعلى (٩٩%) مع بقاء قيمة   الخطأ المسموح به على ماهي عليه،فإن حجم العينة سوف يزداد، حيث يكون:

$$\text{حجم العينة} = \left(\frac{٢,٥٨\ (١٥٠)}{١٠}\right)^{٢} = ١٤٩٧,٦٩$$

أي أن حجم العينة المطلوبة يكون ١٤٩٨ مفردة، وذلك بحدود خطأ ١٠ ومستوى ثقة ٩٩% (٢,٥٨ وحدات الانحراف المعياري)، وهنا قد يجد الباحث أن حجـــم العينـــة كبيراً، وفي هذه الحالة يمكنه تقليل حجم العينة بحيث يكون في حدود التكلفـــة والوقـــت المتاح. إن تقليل العينة في هذه الحالة يكون بزيادة قيمة الخطأ المسموح به، كأن يكـــون ١٥ (بدلاً من ١٠ كما هو في المثال السابق)، بمعنى أنه إذا كان متوسط الإنفـــاق كمـــا ستكشفه العينة يساوي ٣٤٠ جنيها، فإننا نعتبره يتراوح مابين ٣٢٥ إلى ٣٥٥ جنيها ، وهنا يتم تقدير حجم العينة كالآتي:

$$\text{حجم العينة} = \left(\frac{٢,٥٨\ (١٥٠)}{١٥}\right)^{٢} = ٦٦٥,٦٤$$

أي أن حجم العينة المطلوبة هو ٦٦٦ مفردة ، وذلك بخطأ مسموح به قدره ١٥ ومستوى ثقة ٩٩% (٢,٥٨ وحدة انحراف معياري)

وكثيرا ما يكون حجم المجتمع الذي ستسحب منه العينة معروفاً، وفي هذه الحالة يتم استخدام معامل تصحيح إضافي ومعادلته:

$$ع م = \frac{ع١}{١ + \frac{(ع١ - ١)}{ن}}$$

حيث إن :

ع م: حجم العينة المعدلة (بعد استخدام معامل التصحيح الإضافي)

ع ١ : حجم العينة الذي سبق تحديده (وهو في حالتنا هذه ٣٨٤ مفردة)

ن: حجم المجتمع

فالجزء الأسفل (المقام) في المعادلة المذكورة هو: حجم العينة الذي حصلنا عليه مطروحاً منه واحد صحيح مع القسمة على حجم المجتمع، أي: (٣٨٤-١)÷ ن، ويضاف ناتج هذه القسمة إلى واحد صحيح، أي : ١+ (٣٨٤-١) ÷ ن فيكون لدينا قيمة معينة،وبقسمة حجم العينة الأصلي على هذه القيمة نحصل على حجم العينة المعدل (ع م)

فإذا افترضنا أن حجم العينة الذي حصلنا عليه هو ٣٨٤ وكان حجم المجتمع الأصلي هو(٣٧٠٠٠) مفردة، فإن حجم العينة المعدل يكون:

$$ع م = \frac{٣٨٤}{١ + \frac{(٣٨٤ - ١)}{٣٧٠٠٠}} = ٣٨٠,٢$$

أي أن حجم العينة المعدل أو المصحح يساوي ٣٨٠ مفردة

رابعاً: تحديد حجم العينة عند هامش الخطأ:

إن تحديد حجم العينة عند هامش خطأ معين يعتبر مسألة هامة للبحوث العلمية أياً كـــان موضوعها. لنفرض أننا نريد تحديد الحجم المناسب للعينة بهامش خطأ ٢,٥% وذلـــك عند مستوى الثقة ٠,٩٥ في هذه الحالة نفترض أن حجم العينة يساوي ٠,٥ وبـــالطبع سوف ينتج عن ذلك حجم عينة أكبر مما هو مطلوب، لكنه يجعلنا في مأمن لتحديد الحجم المناسب، كما أن هامش الخطأ سوف يقلل من حجم العينة الأكـــبر. وبمـــا أن القيمـــة المعيارية لمستوى الثقة ٠,٩٥ هي ١,٩٦ فإن حجم العينة المناسب عند هـــامش الخطـــأ ٢,٥% يكون بموجب المعادلة:

$$١,٩٦ \sqrt{\frac{٠,٥ \times ٠,٥}{ن}} = ٠,٠٢٥$$

أي أن:

$$\sqrt{\frac{٠,٢٥}{ن}} = \frac{٠,٠٢٥}{١,٩٦}$$

وبالتالي فإن طرفي المعادلة هما:

$$= \frac{٠,٢٥}{ن} = \left(\frac{٠,٠٢٥}{١,٩٦}\right)^{٢}$$

$$ن = \frac{٠,٢٥}{} = ١٥٣٦,٦٤$$

$(\text{١,٩٦} \div \text{٠,٠٢٥})^{٢}$

أي أنه إذا اختيرت عينة قوامها ١٥٣٧ فإن هامش خطأ المعاينة عند مستوى ثقة ٠,٩٥ يكون في حدود ± ٢,٥% (وليس أكثر أو أقل من ذلك)

خامساً: استخدام الجداول الإحصائية:

وضع الإحصائيون جداول لتحديد حجم العينة المناسب ، فهذه الجداول تتضمن حجم العينة الذي يقابل عدد مفردات المجتمع، ومن تلك الجداول نقدم الجدول الآتي، والذي يتضح منه – على سبيل المثال- أنه إذا كان حجم المجتمع عشر مفردات، فإن حجم العينة يكون عشر مفردات (أي أنه يتعين دراسة كل المجتمع)، ويوضح الجدول أيضاً أنه إذا كان حجم المجتمع ٢١٠ فإن حجم العينة المناسب يكون ١٣٦ (أنظر نهاية العمودين الأول والثاني يسار)، أما إذا كان حجم المجتمع ١٢٠٠ فإن العينة تكون ٢٩١ مفردة (بداية العمودين الأول والثاني يمين)، وإذا كان حجم المجتمع ٧٥٠٠٠ فإن العينة تكون ٣٨٢ مفردة، أما إذا كان حجم المجتمع مليون أو أكثر، فإن حجم العينة يمكن أن يكون ٣٨٤ مفردة ...وهكذا:

Population	Sample	Population	Sample	Population	Sample
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	228	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000∝	384

وكما هو واضح من الجدول فإنه لا يتضمن جميع الأرقام التي تعبر عن مفردات المجتمع، على سبيل المثال فإن المجتمع قد يضم ٢٤٥٨ مفردة، وفي هذه الحالة يتم تحديد حجم العينة بأقرب حجم موضح بالجدول، حيث يتضح أن القيمة ٢٤٥٨ تقع بين ٢٤٠٠ و٢٦٠٠ لكنها أقرب إلى ٢٤٠٠، وعندما يكون المجتمع ٢٤٠٠ فإن الجدول يوضح أن حجم العينة يكون ٣٣١ مفردة، لكن إذا افترضنا أن لينا مجتمع عدده ٢٥٨١ فإن أقرب عدد مناظر له بالجدول هو ٢٦٠٠ وبالتالي ـــ حسب الجدول يكون حجم العينة ٣٣٥ مفردة.

المبحث الثاني

إجراءات المعاينة الإحصائية

أولاً: العشوائية في اختيار العينات

إن مصطلح العشوائية Randomness في اختيار العينة لا يعنى الفوضى أو عدم النظـــام كما قد يفهم من لفظ (العشوائية) وإنما يعنى أن جميع أفراد مجتمع البحث تتاح لهم فرص متساوية لأن يتم اختيارهم ضمن العينة. فإتاحة الفرصة المتساوية لجميع المفردات لأن يتم اختيارها ضمن العينة هو الأسلوب الأمثل في اختيار عينة البحـــث العلمـــي، وللمعاينـــة العشوائية مزايا أساسية أولها حذف التحيز Eliminating Bias فالأسلوب العـــشوائي يؤكد أن مفردات العينة تم اختيارها بدون تحيز، وعلى الرغم من أن الاختيار العشوائي لا يضمن أن تكون العينة ممثلة للمجتمع إلا أنه يحذف مخاطر الاختيار المتحيز. أمـــا الميـــزة الثانية للاختيار العشوائي فهي تحديد الثقة Determining Confidence فهذا الاختيـــار يضع الأساس العلمي لتحديد درجة الثقة في الاستنتاج الإحصائي ، هذا الاستنتاج الذي لا يمكن تنفيذه أو الاعتداد به إذا كانت مفردات العينة تم اختيارها بطريقة أخـــرى غـــير الطريقة العشوائية. إن فترة الثقة للنسبة المحتمل وجودها في المجتمع يمكن الاعتداد بها فقط إذا كانت النتائج مستمدة من عينة عشوائية، فإذا افترضنا أن دراسة أجريت على عينـــة عشوائية قوامها ٥٠٠ مفردة، وتبين أن ٤٠٠ يشاهدون التليفزيون، أي أن نسبة المشاهدة هي ٨٠% فإذا أردنا تحديد وجود هذه النسبة في المجتمع بدرجة ثقة ٩٥% مثلاً فـــإن القيمة المعيارية المقابلة لهذه الدرجة هي ١,٩٦ وتكون فترة الثقة :

$$٨٠\ \% \pm \sqrt{\frac{٠,٢٠ \times ٠,٨٠}{٥٠٠}}$$

$$= ٠,٨٠ \pm ٠,٠١٨$$

أي أن فترة الثقة تتراوح ما بين ٠,٧٨٢ إلى ٠,٨١٨

الميزة الثالثة للاختيار العشوائي هي التحكم في خطأ المعاينـــة Controlling Sampling Error فمن خلال الاختيار العشوائي ومعرفة حجم العينة، يمكن تحديـــد المـــستوى المرغوب من خطأ المعاينة (كأن نجعله في أضيق الحدود) ، علماً بأنه من غير الممكن تحقيق مستوى مقبول من خطأ المعاينة إذا تم اختيار مفردات العينة بأسلوب غير عشوائي، ففـــي

المثال فإن خطأ المعاينة هو ٠,٠١٨.

وإذا كان الاحتيار العشوائي هو إتاحة فرص متساوية ومستقلة لكل مفردات المجتمع لأن يتم اختيارها في العينة، فإن ذلك يتطلب تحديد مجتمع الدراسة؛ أي المجتمع الأصلي الذي ستسحب منه العينة، ومعرفة خصائص هذا المجتمع من حيث التوزيع الجغرافي، ومدى التجانس في الخصائص الديموجرافية، تلك الخصائص التي تمثل متغيرات يمكن أن تؤثر في الظاهرة أو الموضوع محل البحث، وقد يكون مجتمع الدراسة عبارة عن التلاميذ في مرحلة دراسية معينة، أو أولياء الأمور، أو الأطفال في سن ما قبل المدرسة، أو ربات البيوت، أو شباب الجامعات، أو الموظفين أو المطلقين، أو الأرامل، أو التلاميذ الموهوبين، أو ذوى الإعاقة ... الخ.

فإذا افترضنا أن هناك فصلاً مدرسياً يتكون من أربعين تلميذاً، ونريد اختيار عينة عشوائية من هؤلاء التلاميذ، على أن تكون العينة ثمانية أفراد مثلاً (ن= ٨)، فإن هذا الاختيار يتم كالآتي:

- إعطاء رقم مسلسل لكل تلميذ، وذلك في قائمة مستقلة تتضمن أرقام وأسماء التلاميذ.
- كتابة كل رقم على ورقة أو بطاقة مستقلة، فيكون لدينا أربعون بطاقة تحمل أرقاماً مسلسلة من (١) إلى (٤٠).
- خلط البطاقات جيداً، واختيار ثمانية بطاقات، فتكون الأرقام المدونة على هذه البطاقات هي الأسماء التي تتضمنها العينة.

فإذا كانت هذه البطاقات تحمل أرقام ٣، ٥، ٢١، ١٣، ٢٠، ١٨، ٢٦، ٣٢ فإن ذلك يعني أن العينة ستجرى على التلاميذ الذين يحملون هذه الأرقام في قائمة الأسماء. وهنا تكون مفردات العينة قد تم سحبها عشوائياً. وعندما تكون قائمة المفردات كبيرة، فإن مفردات العينة يتم سحبها بواسطة الحاسب الآلي من خلال البرامج الإحصائية الجاهزة وقد يتم السحب مع الإعادة بمعنى أننا نختار البطاقة الأولى، ثم ندون ما تحمله من رقم في ورقة مستقلة، ثم تعاد البطاقة مرة أخرى ضمن بقية البطاقات، وهكذا كلما سحبنا بطاقة ندون الرقم الموجود بها ثم نعيدها ثانية ونواصل الاختيار حتى نستكمل العينة، وقد يحدث أن نختار بطاقة سبق اختيارها، وهنا تعاد ثانية ضمن البطاقات التي يتم السحب منها، وتستمر عملية الاختيار على ألا تتضمن بطاقات سبق اختيارها.

أما السحب مع عدم الإعادة، فيعنى عدم إعادة البطاقة التي تم اختيارها إلى بقية البطاقات التي يتم السحب منها.

ثانياً: التكامل بين العشوائية والأساليب الأخرى:

على الرغم من مزايا الأسلوب العشوائي وضرورته في اختيار العينة، إلا أنه لا يضمن أن تكون العينة ممثلة للاختلافات أو التباينات الموجودة في المجتمع الذي سحبت منه هذه العينة.لنفرض أن لدينا فصلاً مدرسياً يضم ٢٠ من الذكور، مقبل ٢٠ من الإناث، فإذا كنا نريد سحب عينة مكونة من ٨ مفردات، فإنه يفترض أن تضم ٤ ذكور مقابل ٤ إناث، غير أن ذلك قد يتحقق وقد لا يتحقق ، فقد تأتي العينة موزعة بالتساوي بين الجنسين، كما قد تأتي متضمنة ٥ ذكور مقابل ٣ إناث، أو ستة إناث مقابل اثنين من الذكور أو غير ذلك من التوزيعات التي تكون معها العينة غير ممثلة للمجتمع من حيث متغير النوع.

أضف إلى ذلك أن مجتمع البحث يمكن أن يكون كبيراً (بالألوف أو بالملايين) وموزعاً على مساحة جغرافية شاسعة، وبالتالي يصعب حصر الأسماء في قوائم متسلسلة (إلا إذا كانت هذه القوائم موجودة لدى الجهات المعنية، ويمكن استخدام الحاسب الآلي في اختيار عينة عشوائية). ولكفاءة الاختيار العشوائي والتغلب على الصعوبات التي تواجهه، يلجأ الباحثون إلى عدة طرق أهمها: استخدام جداول الأرقام العشوائية، تمثيل فئات المجتمع (تقسيم المجتمع إلى طبقات متجانسة ثم سحب مفردات من كل طبقة)، المزج بين العشوائية والانتظام في اختيار مفردات العينة:

(أ) استخدام جداول الأرقام العشوائية:

كثيراً ما يستعين الباحثون بجداول الأرقام العشوائية في اختيار مفردات العينة العشوائية ، وهذه الجداول تم إعدادها وفق منهج علمي وتتضمن أرقاماً تم تنظيمها بطريقة خاصة بحيث يمكن استخدامها في اختيار مفردات العينات دون تحيز، إذ إنها تتيح فرصاً متساوية لجميع المفردات بأن يتم اختيارها ضمن العينة، فإذا أردنا أن نسحب عينة قوامها ٢٠٠ طالب من بين طلاب إحدى الكليات والبالغ إجمالهم ٥٠٠ طالب، فإن المطلوب أن يكون لدينا قائمة مسلسلة بأسماء هؤلاء الطلاب الخمسمائة، ثم نستخدم جدول الأرقام العشوائية بأن نضع أصبعنا بشكل عشوائي على أي رقم في هذا الجدول، وننظر إلى الأعداد الثلاثة الأولى (نظراً لأن حجم العينة المطلوبة ٢٠٠ وهذا الرقم – ٢٠٠- يتكون من ثلاثة أعداد) فإذا كان هذا الرقم الذي وضعنا إصبعنا عليه يقل عن ٥٠٠، فإنه يتم رصده في ورقة مستقلة، أما إذا كان يزيد عن خمسمائة، فيتم تجاوزه والانتقال إلى الرقم الذي يليه، وهكذا تستمر العملية حتى نحصل على ٢٠٠ رقم من جدول الأرقام العشوائية، وبرصد تلك الأرقام في ورقة مستقلة، فإننا نختار أسماء الطلاب الذين يحملون نفس الأرقام في قائمة أسماء الطلاب.على سبيل المثال إذا كانت الأرقام التي تم رصدها من جدول الأرقام

العشوائية هي (١٠٠٩٧)، (٣٧٥٤٢)، (١٢٨٠٧) فإن ذلك يعني اختيار الطلاب أرقام (١٠٠) ، (٣٧٥)، (١٢٨) في قائمة الأسماء ، والجدول التالي نموذج لأحـــد الجـــداول العشوائية التي يمكن الاستعانة بها في الاختيار العشوائي لمفردات العينات التي تقـــل عـــن عشرة آلاف مفردة:

10097	85017	84532	13618	23157	86952	02438	76520
37542	16719	82789	69041	05545	44109	05403	64894
08422	65842	27672	82186	14871	22115	86529	19645
99019	76875	20684	39187	38976	94324	43204	09376
12807	93640	39160	41453	97312	41548	93137	80157
66065	99478	70086	71265	11742	18226	29004	34072
31060	65119	26486	47353	43361	99436	42753	45571
85269	70322	21592	48233	93806	32584	21828	02051
63573	58133	41278	11697	49540	61777	67954	05325
73796	44655	81255	31133	36768	60452	38537	03529
98520	02295	13487	98662	07092	44673	61303	14905
11805	85035	54881	35587	43310	48897	48493	39808
83452	01197	86935	28021	61570	23350	65710	06288
88685	97907	19078	40646	31352	48625	44369	86507
99594	63268	96905	28797	57048	46359	74294	87517
65481	52841	59684	67411	09243	56092	84369	17468
80124	53722	71399	10916	07959	21225	13018	17727
74350	11434	51908	62171	93732	26958	02400	77402
69916	62375	99292	21177	72721	66995	07289	66252
09893	28337	20923	87929	61020	62841	31374	14225
91499	38631	79430	62421	97959	67422	69992	68479
80336	49172	16332	44670	35089	17691	89246	26940
44104	89232	57327	34679	62235	79655	81336	85157
12550	02844	15026	32439	58537	48274	81330	11100
63606	40387	65406	37920	08709	60623	2237	16505
61196	80240	44177	51171	08723	39323	05798	26457
15474	44910	99321	72173	56239	04595	10836	95270
94557	33663	86347	00926	44915	34823	51770	67897
42481	86430	19102	37420	41976	76559	24358	97344
23523	31379	68588	81675	15694	43438	36879	73208
04493	98086	32533	17767	14523	52494	24826	75246
00549	33185	04805	05431	94598	97654	16232	64051
35963	80951	68953	99634	81949	15307	00406	26898
59808	79752	02529	40200	73742	08391	49140	45427
46058	18633	99970	67348	49329	95236	32537	01390
32179	74029	74717	17674	90446	00597	45240	87379
69234	54178	10805	35635	45266	61406	41941	20117
19565	11664	77602	99817	28573	41430	96382	01758
45155	48324	32135	26803	16213	14938	71961	19476
94864	69074	45753	20505	78317	31994	98145	36168

(ب) تمثيل فئات المجتمع:

بمعنى تقسيم المجتمع إلى فئات، تضم كل فئة المفردات التي تشترك في صفة معينة، ومن بين كل فئة يتم السحب العشوائي للمفردات المطلوبة. ولتوضيح هذه الطريقة ببساطة، افترض أن فصلاً دراسياً يضم مائة تلميذ وتلميذة، منهم (٧٠) ذكور، مقابل (٣٠) إناث. إن اختيار عينة من عشرة مفردات مثلاً يعنى أن تتم عملية الاختيار عشوائياً على مستوى الذكور على حدة، ثم على مستوى الإناث، بحيث تتضمن العينة مفردات من الجنسين، والاختيار هنا يمكن أن يتم بالأسلوب العشوائي البسيط، أو بالأسلوب العشوائي المنتظم، والعينة التي يتم اختيارها يمكن توزيعها بالتساوي بين الجنسين، بمعنى خمسة ذكور مقابل خمس إناث، كما يمكن توزيعها بالتناسب بين الجنسين، أي أن يكون عدد المفردات التي يتم اختيارها من كل جنس تتناسب مع إجمالي عدده الأصلي، وهنا فإن العينة التي تضم عشر مفردات تعني أن يتم سحب سبع مفردات من الذكور، وثلاث مفردات من الإناث.

إن هذا المثال لمجرد التبسيط، لأن تقسيم التلاميذ إلى فئات، يمكن أن يتم على أساس مستوى التحصيل، الذكاء، السن، منطقة الإقامة ... الخ، كما أن المجتمع قد يكون بالآلاف أو حتى بالملايين.افترض على سبيل المثال، أن إجمالي عدد السكان في مدينة معينة هو ١٣٩٦٨١، منهم ٩٦٢٤١ من الذكور، مقابل ٤٣٤٤٠ من الإناث، أي أن السكان يتوزعون بين الذكور بنسبة ٦٩%، والإناث بنسبة ٣١%، وأننا نريد سحب عينة عشوائية حجمها (٥٠٠) مفردة من الجنسين لإجراء دراسة عن قراءة الصحف، في هذه الحالة، فإن عدد المفردات التي يتعين سحبها من فئة الذكور يكون كالآتي:

$$\frac{\text{عدد الذكور}}{\text{إجمالي السكان}} \times \text{حجم العينة}$$

$$= \frac{٩٦٢٤١}{١٣٩٦٨١} \times ٥٠٠ = ٣٤٤,٥$$

أي (٣٤٥) مفردة. وبالتالي، فإن المفردات التي يتعين سحبها من الإناث يساوى:

٥٠٠ – ٣٤٥ = ١٥٥ مفردة.

ويمكن الحصول على عدد الإناث في العينة كالآتي:

$$\frac{\text{عدد الإناث}}{\text{إجمالي السكان}} \times \text{حجم العينة}$$

$$= \frac{٤٣٤٤٠}{١٣٩٦٨١} \times ٥٠٠ = ١٥٥ \text{ مفردة تقريباً}$$

أي أن العينة تتوزع بين الذكور بنسبة ٦٩٪، والإناث بنسبة ٣١٪، وهى نسبة كل جنس في مجتمع البحث، فكأن تركيب العينة عددياً يعكس تركيب المجتمع من حيث النوع.

وإذا كان هذا يسمى بالتوزيع المتناسب، فإن هناك نوعاً آخر من التوزيع الطبقي للعينة يسمى التوزيع الأمثل، وهو يمكن الحصول عليه فقط عندما يكون الانحراف المعياري لكل فئة أو طبقة معروفاً. افترض أننا نريد سحب عينة قوامها ٤٠٠ مفردة من مجتمع يضم ثلاث مناطق (منطقة ريفية، منطقة حضرية، منطقة بدوية) ، وذلك لدراسة استخدام القنوات الفضائية الأجنبية في تلك المناطق، وكان سكان كل منطقة كالآتي:

- المنطقة الريفية ١٨٥١٣٢١
- المنطقة الحضرية ١٦٢٣٤١٠
- المنطقة البدوية ١٤١١٢٧٣

أي أن إجمالي سكان هذه المناطق الثلاث هو(٤٨٨٦٠٠٤) وقد تم اختيار ١٠٠ مفردة، وأجريت عليهم دراسة استطلاعية من سؤال واحد- بشأن عدد الساعات التي يقضيها الشخص أسبوعيا في مشاهدة القنوات الفضائية الأجنبية، وقد كشف تحليل استجابات المفحوصين عن أن الانحرافات المعيارية لمتوسط ساعات المشاهدة الأسبوعية لكل مجموعة كالآتي:

- المنطقة الريفية: الانحراف المعياري= ٢,٦
- المنطقة الحضرية الانحراف المعياري= ١,٧
- المنطقة البدوية : الانحراف المعياري= ٣,٤

فإذا كان المطلوب هو سحب عينة قوامها ٤٠٠ مفردة ، فإن العدد الأمثل الذي يتعين سحبه من كل مجموعة يمكن الحصول عليه بموجب المعادلة:

$$\frac{\text{عدد المجموعة المعنية} \times \text{انحرافها المعياري}}{\text{عدد مفردات كل مجموعة} \times \text{انحرافها المعياري}} \times \text{حجم العينة}$$

أي أن المقام يظل ثابتاً (فالمقام عبارة عن مجموع عدد مفردات كل مجموعة من المجموعات الثلاث مضروباً في انحرافها المعياري))، أما البسط فهو عبارة عن عدد مفردات المجموعة المطلوب تحديد عدد مفرداتها مضروباً في الانحراف المعياري لتلك المجموعة، وباستخراج قيمة المقام (الثابت) فإن هذه القيمة تساوي:

(١٨٥١٣٢١×٢,٦) + (١٦٢٣٤١٠ × ١,٧)

+ (١٤١١٢٧٣ ×٣,٤) = ١٢٣٧١٥٥٩,٨

أي أن قيمة المقام أو الثابت هي ١٢٣٧١٥٥٩,٨

بناء على ذلك، فإن عدد المفردات التي يتم سحبها من مجموعة الريف=

$$\frac{١٨٥١٣٢١ \times ٢,٦}{١٢٣٧١٥٥٩,٨} \times ٤٠٠ = ١٥٦$$

بينما يكون عدد المفردات التي يتعين سحبها من مجموعة الحضر=

$$\frac{١٦٢٣٤١٠ \times ١,٧}{١٢٣٧١٥٥٩,٨} \times ٤٠٠ = ٨٩$$

أما عدد المفردات التي يتعين سحبها من مجموعة البدو فهو:

$$\frac{٣,٤ \times ١٤١١٢٧٣}{١٢٣٧١٥٥٩,٨} \times ٤٠٠ = ١٥٥$$

أي أنه حسب التوزيع الأمثل، فإن العينة تتوزع بين مجموعة الريف بواقع ١٥٦ مفردة، ومجموعة الحضر بواقع ٨٩ مفردة، ومجموعة البدو بواقع ١٥٥ مفردة ، وهذا يختلف عما إذا كنا قد اتبعنا طريقة التوزيع المتناسب.

فحسب التوزيع المتناسب، نجد أن مجموعة الريف تضم ١٨٥١٣٢١ أي أنها تشكل ٣٨% تقريباً من مجمل المناطق. فإذا كان حجم العينة الكلية ٤٠٠ مفردة، فإن التوزيع المتناسب يعني أن يكون عدد المفردات المسحوبة من مجموعة الريف هو ١٥٢ (بمعنى ٤٠٠ × ٣٨% = ١٥٢) ، أما فيما يخص مجموعة الحضر، فإنها تشكل ٣٣,٢% من مجمل مناطق البحث، وبالتالي فإن التوزيع المتناسب يقضي بأن يكون عدد مفرداتها في العينة هو١٣٣ (بمعنى ٤٠٠× ٣٣,٢ %) وعلى مستوى مجموعة البدو، فإنها تشكل ٢٨,٩% من مجمل مجتمع البحث، وبالتالي فإن التوزيع المتناسب يقضي بأن يكون عدد مفرداتها في العينة يساوي١١٥ (بمعنى ٤٠٠× ٢٨,٩%) ، هكذا يتضح أنه عند مقارنة توزيع مفردات العينة حسب أسلوب التوزيع الأمثل وأسلوب التوزيع المتناسب، فإن هناك اختلافاً يوضحه الجدول الآتي:

المجموعة	مجمل المجتمع	حسب التوزيع المتناسب	حسب التوزيع الأمثل
مجموعة الريف	١٨٥١٣٢١	١٥٢	١٥٦
مجموعة الحضر	١٦٢٣٤١٠	١٣٣	٨٩
مجموعة البدو	١٤١١٢٧٣	١١٥	١٥٥
الإجمالي	٤٨٨٦٠٠٤	٤٠٠	٤٠٠

من الواضح أنه حسب التوزيع المتناسب، فإن نسبة مفردات كل مجموعة في العينة تتساوى مع نسبة عدد هذه المجموعة في المجتمع، على سبيل المثال، فإن مجموعة الريف تضم

١٨٥١٣٢١ أي أنها تشكل قرابة ٣٨% من الإجمالي البالغ ٤٨٨٦٠٠٤ وفي الوقت نفسه نجد أن عدد مفردات مجموعة الريف في العينة هو ١٥٢ مفردة أي ما يعادل ٣٨% من العينة، المنطق نفسه مع اختلاف الأرقام فيما يخص مجموعتي الحضر والبدو. لكن عند استخدام أسلوب التوزيع الأمثل اختلف الأمر بعض الشيء فقد انخفض عدد مفردات مجموعة الحضر في العينة ليصبح ٨٩ مفردة فقط (بعد أن كان ١٣٣ مفردة حسب التوزيع المتناسب) ، هذا الانخفاض يفسر بأن مجموعة الحضر هي أكثر المجموعات تجانساً، حيث تنخفض قيمة الانحراف المعياري إلى ١,٧ (ومن المعروف أنه يمكن تقليل حجم العينة كلما كان المجتمع أكثر تجانساً) أما مجموعة البدو، فإنها أكثر مجموعات العينة تبايناً حيث ترتفع قيمة الانحراف المعياري إلى ٣,٤ ومن هنا ارتفع عدد مفرداتها من ١١٥ مفردة حسب أسلوب التوزيع المتناسب إلى ١٥٥ مفردة حسب التوزيع الأمثل (ومن المعروف أنه يتعين زيادة حجم العينة كلما المجتمع أكثر تبايناً) ، أما مجموعة الريف، فقد كان عدد مفرداتها في العينة١٥٢ وذلك حسب التوزيع المتناسب، ارتفع إلى ١٥٦ حسب التوزيع الأمثل، وذلك يعزى أيضاً إلى التباين، فإذا كانت مجموعة الريف أقل تبايناً عن مجموعة البدو، إلا أنها أكثر تبايناً من مجموعة الحضر. بوجه عام، فإن التوزيع الأمثل هو الأشد دقة عند اختيار العينات لأنه يأخذ تباين المجتمع في الاعتبار، لكن المشكلة في تطبيق هذا الأسلوب تتمثل في ضرورة المعرفة المسبقة بقيمة الانحراف المعياري لكل مجموعة أو فئة أو طبقة، وهذه المعرفة نادراً ما تتحقق خاصة عندما يكون عدد المجموعات كبيراً، وعلى الرغم من إمكانية تقدير الانحراف المعياري للمجموعات بشأن المتغيرات التي يتم على أساسها اختيار العينة، إلا أن هذا التقدير عرضة للخطأ في كثير من الأحيان، وعادة يستخدم الباحثون التوزيع المتناسب، أو المتساوي إلا في الحالات التي يكون الانحراف المعياري مقدراً وفق معطيات قوية

(ج) انتظام الاختيار

يقوم هذا الأسلوب على العشوائية مضافاً إليه الانتظام في اختيار مفردات العينة، وبينما تكون العشوائية في اختيار المفردة الأولى، فإن الانتظام يكون في تساوي المدى Range بين المفردات التي يتم اختيارها في العينة. لتوضيح ذلك، افترض أننا نريد سحب عينة تتكون من ثمان مفردات من فصل مدرسي يضم أربعين تلميذاً

في هذه الحالة نعطى للتلاميذ أرقاماً متسلسلة في القائمة المتضمنة أسماؤهم، فيكون لدينا أربعين رقماً ، من ١ إلى ٤٠، ثم نحدد المدى Range بقسمة إجمالي عدد تلاميذ الفصل على حجم العينة المطلوبة، أي:

$$\frac{٤٠}{٨} = ٥$$

أي أن قيمة المدى تساوي خمسة، بعد ذلك نضع أرقاماً مسلسلة من ١ إلى ٥، بحيث يكتب كل رقم في ورقة مستقلة، وتخلط هذه الأوراق الخمس، ويتم اختيار إحدى الأوراق منها. افترض أننا اخترنا الورقة التي تحمل الرقم ٣، فيكون التلميذ الأول في العينة هو التلميذ رقم ٣ بالقائمة الأصلية (المتضمنة أسماء جميع التلاميذ بأرقام متسلسلة) ثم يضاف الرقم ٥، فيكون التلميذ الثاني هو ٨، وهكذا إلى أن نصل إلى العدد المطلوب اختياره في العينة، وبالتالي تضم العينة التلاميذ أرقام: ٣، ٨، ١٣، ١٨، ٢٣، ٢٨، ٣٣، ٣٨ (لاحظ أن بين كل مفردة أخرى عدد متساو من المفردات، هذا العدد المتساوي هو المدى، ويبلغ في مثالنا هذا خمس مفردات). وكما هو واضح، فإن مفردات العينة تتحدد بمجرد اختيار المفردة الأولى، وهذا يتطلب وجود قائمة شاملة وبأرقام متسلسلة لجميع مفردات المجتمع (وهذا لا يتحقق في أحيان كثيرة)، كما أنه من الضروري وجود قائمة تضم جميع أفراد المجتمع، وهذه القائمة مرتبة عشوائياً (كأن تكون الأسماء مدونة في تلك القائمة بصرف النظر عن الجنس (ذكور & إناث)، أو الترتيب الأبجدي ...الخ)، فإذا لم تكن المفردات في القائمة الأصلية مرتبة عشوائياً، فإن ذلك قد يؤدي إلى استبعاد بعض المجموعات بحيث لا يتم اختيارها ضمن العينة رغم أنها موجودة في المجتمع

المبحث الثالث

خطأ المعاينة

إن علاقة العينة بالمجتمع الذي سحبت منه تعتمد على حجم العينة وطريقة اختيارها. العينة الصحيحة هي التي تمثل المجتمع الذي سحبت منه تمثيلاً صادقاً، وتقترب العينة من الأصل الذي سحبت منه كلما اقتربت مقاييسها الإحصائية من مقاييس ذلك الأصل، على سبيل المثال، نفرض أن لدينا مجتمعاً يضم يضم ٣٠ مليون شخص تتراوح أعمارهم ما بين ٢٠ سنة إلى ٩٠ سنة، وكان متوسط العمر لهذا المجتمع هو ٦٥ سنة، ومن هذا المجتمع سحبنا عينة عشوائية قوامها ٤٠٠ شخص، وحسبنا متوسط السن لهذه العينة، وباستخدام الطرق الإحصائية يمكن المقارنة بين متوسط الأعمار في العينة ومتوسط الأعمار في المجتمع، فإذا تبين أنه لا توجد فروق جوهرية بين المتوسطين، فإن هذا يعني أن متوسط أعمار العينة لا يختلف اختلافاً جوهرياً عن متوسط الأعمار في المجتمع الذي سحبت منه، وفي

هذه الحالة تكون العينة ممثلة للمجتمع تمثيلاً صادقاً من حيث السن (ومن الضروري هنا أخذ متغيرات أخرى بالاعتبار مثل الجنس ومنطقة الإقامة..الخ)
وعلى الرغم من ذلك، فإن العينة مهما كانت دقة اختيارها عشوائياً لا تكون ممثلة للمجتمع مائة بالمائة، وذلك بسبب ما يعرف بخطأ المعاينة Sampling Error والذي ينعكس على المقاييس الإحصائية التي نحصل عليها جراء المعالجة الإحصائية لبيانات العينة بحيث تأتي مختلفة بعض الشيء عن المجتمع. إن ذلك ينطبق على كافة المقاييس الإحصائية، وسوف نوضح هنا مثالين فقط هما الخطأ المعياري للمتوسط والخطأ المعياري للنسبة.

أولاً: الخطأ المعياري للمتوسط:

يتم حساب الخطأ المعياري للمتوسط الحسابي للعينة بمعلومية الانحراف المعياري وحجم العينة وذلك بموجب القانون:

$$\frac{ع}{\sqrt{ن}}$$

فإذا كان لدينا عينة من ٤٠٠ شخص، وطبقنا عليها مقياس استخدام وسائل الإعلام، وتبين أن متوسط درجة العينة على هذا المقياس ٨,٧ درجة بانحراف معياري هو ٣,٦ فإن الخطأ المعياري للمتوسط (ع م) يكون:

$$ع\ م = \frac{٣,٦}{\sqrt{٤٠٠}} = ٠,١٨$$

فإذا كانت قيمة الخطأ المعياري لمتوسط درجة العينة هي ٠,١٨ فإن حدود هذا المتوسط هي:

الحد الأدنى = المتوسط - الخطأ المعياري = ٨,٧ - ٠,١٨ = ٨,٥٢

الحد الأعلى = المتوسط + الخطأ المعياري = ٨,٧ + ٠,١٨ = ٨,٨٨

أي أن القيمة العددية لمتوسط درجة العينة تمتد من ٨,٥٢ إلى ٨,٨٨

وبما أن التوزيع التكراري للمتوسطات يميل إلى أن يكون اعتدالياً، وبما أن المساحة المحصورة بين ــع و +ع تساوي ٦٨% فإن المساحة المتبقية تساوي ٣٢% أي أن نسبة المساحة المحصورة إلى نسبة المساحة المتبقية هي ٢ إلى ١ هنا، يمكن تقرير أن نسبة

احتمال وجود متوسط درجة المجتمع الذي سحبت منه العينة المذكورة يقع في هذا المدى (من ٨,٢٥ إلى ٨,٨٨) بدرجة ثقة ٢ ودرجة شك ١ غير أنه يمكننا زيادة درجة الثقة إلى ٩٥% وإنقاص درجة الشك إلى ٥% إذا ضربنا الخطأ المعياري في ١,٩٦ (لأن المساحة المعيارية المحصورة بين -١,٩٦ إلى+ ١,٩٦ انحراف معياري تساوي ٠,٩٥ من المساحة الكلية للمنحنى الاعتدالي المعياري)

وبما أن قيمة متوسط درجة العينة تساوي ٨,٧ وبما أن الخطأ المعياري يساوي ٠,١٨ فإن مدى متوسط درجة المجتمع الذي سحبت منه العينة المذكورة يساوي:

٨,٧ ± (٠,١٨ ×١,٩٦)

= ٨,٧ ± ٠,٣٥٢٨

أي أن المدى الذي يقع فيه متوسط المجتمع الذي سحبت من العينة المذكورة يمتد من ٨,٣٤٧٢ إلى ٩,٠٥٢٨

أما إذا استخدمنا درجة الثقة ٠,٩٩ فإن هذا لمدى يكون

٨,٧ ± (٠,١٨× ٢,٥٨) أي:

٨,٧ ± ٠,٤٦٤٤

أي أن المدى الذي يقع فيه متوسط المجتمع الذي سحبت من العينة المذكورة يمتد من ٨,٢٣٥٦ إلى ٩,١٦٤٤ وذلك بمستوى ثقة ٠,٩٩

ثانياً: الخطأ المعياري للنسبة:

يمكن حساب الخطأ المعياري للنسبة التي حصلنا عليها من عينة عشوائية وذلك بهدف تقدير المدى الصحيح الذي تقع فيه تلك النسبة، فإذا كان لدينا عينة عشوائية قوامها ٥٠٠ مفردة، وتبين أن ٣٧% من العينة يقرؤون الصحف اليومية، فإن لهذه النسبة خطأ معين، ويمكن حساب هذا الخطأ (الخطأ المعياري للنسبة) بموجب المعادلة:

$$\sqrt{\frac{\text{النسبة (١– النسبة)}}{\text{ن}}}$$

حيث (ن) هي عدد مفردات العينة ، وبتطبيق هذه المعادلة على المثال المذكور يكون:

$$\sqrt{\frac{٠,٣٧ (١- ٠,٣٧)}{٥٠٠}}$$

أي

$$\sqrt{\frac{٠,٣٧ \times ٠,٦٣}{٥٠٠}} = ٠,٠٢$$

أي أنه إذا كانت نسبة قراءة الصحف اليومية من واقع بحث العينة هي ٣٧% فإن الخطأ المعياري لهذه النسبة هو ٢% وعلى ذلك فإن النسبة المصححة تتراوح مابين ٣٥% إلى ٣٩%

وبما أن النسبة المستمدة من العينة هي ٣٧% فإنها تقع بين النسبتين الأدنى والأعلى ، وبناءً على ذلك يمكننا تقدير نسبة قراءة الصحف اليومية في المجتمع سواء بمستوى ثقة ٠٩٥ أو ٠,٩٩

فإذا اتخذنا مستوى الثقة ٠,٩٥ فإن تقدير النسبة في المجتمع يكون:

$$٠,٠٢ \times ١,٩٦ = ٠,٠٤$$

وحيث إن النسبة المستمدة من العينة هي ٣٧% فإن الحد الأعلى للنسبة في المجتمع هو ٠,٣٧ + ٠,٠٤ = ٤١%

ويكون الحد الأدنى للنسبة المجتمع هو ٠,٣٧ - ٠,٠٤ = ٣٣%

أي أن نسبة قراءة الصحف في المجتمع تقدر ما بين ٣٣% إلى ٤١% وذلك بمستوى ثقة ٠,٩٥

أما إذا ا اتخذنا مستوى الثقة ٠,٩٩ فإن تقدير النسبة في المجتمع يكون:

$$٠,٠٤ \times ٢,٥٨ = ٠,١٠٣٢$$

ويكون الحد الأعلى للنسبة في المجتمع هو ٠,٣٧ + ٠,١٠٣٢ = ٤٧%

ويكون الحد الأدنى للنسبة في المجتمع هو ٠,٣٧ - ٠,١٠٣٢ = ٢٧%

أي أن نسبة قراءة الصحف في المجتمع تتراوح ما بين ٢٧% إلى ٤٧% وذلك بمستوى ثقة ٠,٩٩

مصادر ومراجع:

(أ) مصادر ومراجع عربية:

- بشير صالح الرشيدي (٢٠٠٠) مناهج البحث التربوي: رؤية تطبيقية مبسطة (القاهرة: دار الكتاب الحديث)
- رجاء محمود أبو علام (١٩٩٩) مناهج البحث في العلوم النفسية والتربوية (القاهرة: دار النشر للجامعات)
- زكريا الشربيني (١٩٩٠) الإحصاء اللابارامتري في العلوم النفسية والاجتماعية والتربوية (القاهرة: مكتبة الأنجلو المصرية)
- سعدي شاكر حمودي (٢٠٠٠) علم الإحصاء وتطبيقاته في المجالين التربوي والاجتماعي (عمان: مكتبة دار الثقافة للنشر والتوزيع)
- صلاح الدين محمود علام (٢٠٠٠) تحليل بيانات البحوث النفسية والاجتماعية والتربوية (القاهرة: دار الفكر العربي)
- عاطف أحمد منصور(١٩٩٢) الرياضيات المسلية (القاهرة: مكتبة ابن سينا)
- عبد الحميد محمد نجم & محمد عبد الهادي المحميد (١٩٩٠) الإحصاء الوصفي والتحليلي مع استخدام البرامج الجاهزة (الكويت، د. ن.)
- فؤاد أبو حطب & آمال صادق (١٩٩١) مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والاجتماعية والتربوية (القاهرة: مكتبة الأنجلو المصرية)
- فؤاد البهي السيد (١٩٧٨) علم النفس الإحصائي (القاهرة: دار الفكر العربي)

(ب) مصادر ومراجع أجنبية:

Bekoff,-Marc; Allen,-Colin; Grant,-Michael-C. (1999) Feeding decisions by Steller's jays (Cyanocitta stelleri): The utility of a logistic regression model for analyses of where, what and with whom to eat. ETHOLOGY. Vol 105(5): 393-406

Chen,-Wen-Hung; Thissen,-David (1999) Estimation of item parameters for the three-parameter logistic model using the marginal likelihood of summed scores. British-Journal-of-Mathematical-and-Statistical-Psychology.Vol 52(1): 19-37

Degenholtz,-Howard-B.; Kane,-Rosalie-A.; Kane,-Robert-L.; Finch,-Michael-D. (1999) Long-term care case managers' out-of-home placement decision: An application of hierarchical logistic regression. Research-on-Aging. Vol 21(2): 240-274

Fan,-Xitao; Wang,-Lin (1999) Comparing linear discriminant function with logistic regression for the two-group classification problem. Journal-of-Experimental-Education. 1999 Spr; Vol 67(3): 265-286

Robins,-Garry; Pattison,-Philippa; Wasserman,-Stanley (1999) Logit models and logistic regressions for social networks: III. Valued relations. Psychometrika. Vol 64(3): 371-394

SPSS manuals. Different versions. London.